SIGNALS AND COMMUNICATION TECHNOLOGY

For further volumes:
http://www.springer.com/series/4748

Signals and Communication Technology

Passive Eye Monitoring
Algorithms, Applications and Experiments
R.I. Hammoud (Ed.) ISBN 978-3-540-75411-4

Multimodal User Interfaces
From Signals to Interaction
D. Tzovaras ISBN 978-3-540-78344-2

Human Factors and Voice Interactive Systems
D. Gardner-Bonneau, H.E. Blanchard (Eds.)
ISBN 978-0-387-25482-1

Wireless Communications
2007 CNIT Thyrrenian Symposium
S. Pupolin (Ed.) ISBN 978-0-387-73824-6

Satellite Communications and Navigation Systems
E. Del Re, M. Ruggieri (Eds.)
ISBN 978-0-387-47522-6

Digital Signal Processing
An Experimental Approach
S. Engelberg ISBN 978-1-84800-118-3

Digital Video and Audio Broadcasting Technology
A Practical Engineering Guide
W. Fischer ISBN 978-3-540-76357-4

Three-Dimensional Television
Capture, Transmission, Display
H.M. Ozaktas, L. Onural (Eds.)
ISBN 978-3-540-72531-2

Foundations and Applications of Sensor Management
A.O. Hero, D. Castañón, D. Cochran,
K. Kastella (Eds.) ISBN 978-0-387-27892-6

Digital Signal Processing with Field Programmable Gate Arrays
U. Meyer-Baese ISBN 978-3-540-72612-8

Adaptive Nonlinear System Identification
The Volterra and Wiener Model Approaches
T. Ogunfunmi ISBN 978-0-387-26328-1

Continuous-Time Systems
Y.S. Shmaliy ISBN 978-1-4020-6271-1

Blind Speech Separation
S. Makino, T.-W. Lee, H. Sawada (Eds.)
ISBN 978-1-4020-6478-4

Cognitive Radio, Software Defined Radio, and Adaptive Wireless Systems
H. Arslan (Ed.) ISBN 978-1-4020-5541-6

Wireless Network Security
Y. Xiao, D.-Z. Du, X. Shen
ISBN 978-0-387-28040-0

Terrestrial Trunked Radio – TETRA
A Global Security Tool
P. Stavroulakis ISBN 978-3-540-71190-2

Multirate Statistical Signal Processing
O.S. Jahromi ISBN 978-1-4020-5316-0

Wireless Ad Hoc and Sensor Networks
A Cross-Layer Design Perspective
R. Jurdak ISBN 978-0-387-39022-2

Positive Trigonometric Polynomials and Signal Processing Applications
B. Dumitrescu ISBN 978-1-4020-5124-1

Face Biometrics for Personal Identification
Multi-Sensory Multi-Modal Systems
R.I. Hammoud, B.R. Abidi, M.A. Abidi (Eds.)
ISBN 978-3-540-49344-0

Cryptographic Algorithms on Reconfigurable Hardware
F. Rodríguez-Henríquez
ISBN 978-0-387-33883-5

Ad-Hoc Networking Towards Seamless Communications
L. Gavrilovska ISBN 978-1-4020-5065-7

Multimedia Database Retrieval
A Human-Centered Approach
P. Muneesawang, L. Guan
ISBN 978-0-387-25627-6

Broadband Fixed Wireless Access
A System Perspective
M. Engels, F. Petre
ISBN 978-0-387-33956-6

Acoustic MIMO Signal Processing
Y. Huang, J. Benesty, J. Chen
ISBN 978-3-540-37630-9

Algorithmic Information Theory
Mathematics of Digital Information Processing
P. Seibt ISBN 978-3-540-33218-3

Continuous-Time Signals
Y.S. Shmaliy ISBN 978-1-4020-4817-3

Interactive Video
Algorithms and Technologies
R.I. Hammoud (Ed.) ISBN 978-3-540-33214-5

Handover in DVB-H
Investigation and Analysis
X. Yang ISBN 978-3-540-78629-0

Building Automation Communication Systems with EIB/KNX, LON and BACnet
H. Merz, T. Hansemann, C. Hübner
ISBN 978-3-540-88828-4

Sharing RF Spectrum with Commodity Wireless Technologies
Theory and Practice
J. Kruys, L. Qian ISBN 978-94-007-1584-4

Jan Kruys • Luke Qian

Sharing RF Spectrum with Commodity Wireless Technologies

Theory and Practice

Springer

Jan Kruys
SpectrumConsult
Woerden
5 Tiendhof, 3481GT
The Netherlands
kruys@spectrumconsult.net

Luke Qian
WiTuners, Ltd
8870 Darrow Road, STE F106-123
Twinsburg, OH 44087
USA
luke.qian@wituners.com

ISSN 1860-4862
ISBN 978-94-007-1584-4 e-ISBN 978-94-007-1585-1
DOI 10.1007/978-94-007-1585-1
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2011933604

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This book was born from the perception that there is much more to spectrum use and sharing than one sees reflected in publications, whether academic, commercial or political. The former – in good research style – tend towards reductionism and concentrate on specific, detailed aspects. Commercial publications tend to emphasize the positive aspects and they tend to put promise above practice. Given the ever increasing pace of technology development and recent successes of new wireless technologies, some pundits predict large-scale spectrum scarcity, potentially leading to economic catastrophe. Although economic theory has a hard time explaining recent events that shook the world economy, the notion of spectrum scarcity is intuitively acceptable, even if not correct or immediately relevant.

Commodity wireless technologies have a vast impact on our lives and their growth seems assured. Therefore, the issue of spectrum sharing is important and so is the understanding of the basic facts that shape its possibilities as well as its limitations. Internet guru David Reed has argued that there is no such thing as spectrum scarcity – just like the color green can't be scarce. Of course, he is right – but spectrum sharing is a matter of separating wanted "green" from unwanted "green." The possibilities of doing that are delimited by technology as well as by physical factors: the Shannon-Hartley theorem shows that a receiver is limited in its ability to separate green from green.

Since reality is complicated, theory fragmented and punditry rife, we considered it worthwhile to put our joint experience to work and try to bring together material on the regulatory and technical aspects of spectrum sharing. Given that we see the world at wavelengths that are thousands of times shorter than the wavelengths of the RF devices we use, the complexity of the RF reality is not easily grasped or intuitively clear. We have tried to bring out the importance of the RF environment in shaping the technical possibilities of sharing spectrum. There exists a wealth of research papers on various aspects of spectrum sharing and, therefore, we avoided repeating such work and we have avoided repeating textbook material. To the contrary, we hope that our emphasis on the practical aspects complements textbook material and helps to put the complex mathematics in perspective. Therefore, in putting an overall picture of spectrum sharing together, we cover the why – the

regulatory aspects, as well as the how – the technical aspects. We support the latter with a theoretical taxonomy of dimensions, modes and means of spectrum sharing, and we discuss a number of practical cases of spectrum sharing, using established technologies as examples.

Spectrum sharing affects people in many roles and responsibilities and, therefore, we have tried to serve a broad audience in putting this book together. In the process, we consider spectrum sharing at many levels – from the regulatory level that is largely technology agnostic down to the network operations level that is very much technology specific since it deals with optimizing the capacity of spectrum shared among the network's elements.

The picture that emerges is one of diversity in all aspects, a picture that we hope will stimulate as well as caution.

Acknowledgements

The authors are indebted to the many colleagues they worked with throughout their careers at major communications companies, and they are indebted to many peers in spectrum regulatory fora and standardization committees, notably the IEEE and ETSI. The combination of encouraging perspectives, deep technical knowledge, and a willingness to share what we have encountered so often has helped us greatly in expanding our knowledge and insights.

Special mention deserve our reviewers, Dr. Peter Anker of Delft University, professors Ed Knightly and Stanislav Miskovic at Rice University, and Dr. Stefan Mangold at Disney Research at ETH Zurich. They provided invaluable comments and suggestions to improve this book.

Contents

Part II The Practice of Spectrum Sharing

About the Authors

The authors, Dr. Luke Qian and Mr. Jan Kruys, have an extensive background in wireless technologies that goes back more than 20 years. They have worked together in the subject area for more than 6 years and have been granted a number of patents. They have been deeply involved in product development, as well as in the development of the standards and regulations that have shaped much of the wireless commodity landscape. Both are members of the IEEE Communications Society and contributing authors of the book “Emerging Technologies in Wireless LANs” by Cambridge University Press.

Introduction

For more than a century, wireless technology has been an indispensable part of our lives. It has become a cornerstone of society's communications infrastructure. Its medium – the spatial propagation of electromagnetic fields of a certain frequency or frequency range – is an open, unlimited medium that is accessible to all and sundry. Practical constraints limit the useful range of such frequencies, also referred to as "spectrum", and therefore some form of sharing that useful spectrum is necessary. As the scale of deployment and density of use increase, the development of wireless technologies proceeds apace with advances in signal processing theory, as well as digital signal processing power. Spectrum sharing, in one way or another, becomes more and more feasible.

This book is about the theory and practice of spectrum sharing. Both cellular communications and consumer-oriented communications, like Wi-Fi and Bluetooth, have led to the creation of large new markets for "wireless products" and services. This development has caused the concern that spectrum scarcity could stunt the growth of wireless technologies and services. The response of many spectrum regulators throughout the world in addressing these concerns has been to announce policies that foresee the release of more spectrum for license exempt use or for shared use. An example is the spectrum that is released by the transition to digital TV: the frequencies freed up – sometimes referred to as the "Digital Dividend" – are destined, in part, for new applications that could be license exempt. These policies have attracted many proposals, including some that argue for removing all spectrum regulations, based on the argument that technology – notably cognitive radio – and "the market" will find better ways to distribute and use the available spectrum. However, the theory of spectrum sharing is underdeveloped and practice proves resistive of quick solutions. A case in point is wireless LAN/radar spectrum sharing in the 5 GHz range: seven years after the ITU-R designated this as shared spectrum, the rules for sharing as well as the means for verifying compliance with these rules are not fully mature. Another recent development is the interest in spectrum pricing and trading, which tend to focus on the economic aspects of spectrum sharing rather than at the practical limitations of spectrum sharing. Although the regulatory framework for spectrum trading and sharing is different from the framework that guides

the designation of license exempt spectrum, many of the technical issues are the same or very similar.

Efficient use of shared spectrum imposes severe technical challenges. The ever growing volume of multimedia streams over the Internet pose increasing capacity demands on wireless access networks – whether licensed cellular or license exempt. The cellular network industry is moving towards femtocells as a means to increase their capacity to serve subscribers, and these femtocells have to share frequencies with established macrocell base stations. In the wireless LANs world, overlapping of service areas is common, especially in large-scale, dense deployments in office buildings, on campuses, and in hospitals. In all cases, effective use and management of shared spectrum becomes one of the top issues that the wireless industry is facing nowadays.

The common element in these two examples is that both cellular systems and wireless LANs rely heavily on commodity wireless technologies. This book addresses RF spectrum in the contexts of such technologies; it combines regulatory context, theoretical analysis, and practice-based assessments of RF spectrum sharing from the viewpoint of wireless commodity technologies, in particular that of Wi-Fi, the commercial label of products based on the IEEE 802.11 standard. Although we discuss such subjects as signal transmission, signal propagation, and medium access protocols, the purpose is not to repeat textbook knowledge of basic radio technology, but to point to factors that affect spectrum sharing — positively or negatively.

This book is aimed at a wide audience. Although it addresses basic technical issues, the emphasis is on practice rather than theory. Therefore, it should prove accessible to wireless systems designers, RF engineers, and protocol developers, as well as to planners for large-scale wireless LAN deployments. The combination of theoretical explanation and practice-based analysis is expected to prove useful to/for radio spectrum managers and regulators as well as standards development organizations. In higher education, this book could provide a useful complement to the current curriculum.

To the best of our knowledge, the concepts and ideas described in this book are non-proprietary and can be found in the public domain, including but not limited to Internet-based sources like published Radio Regulatory documents, published industry standards, patents published by the US Patent Office, Wikipedia, and other forms of publications such as trade journals and magazines.

Part I
RF Spectrum Sharing: Background and Theory

There is much discussion of the spectrum sharing and spectrum scarcity, frequently pointing to regulatory authorities as the culprits. However, little attention is being paid to the physical and technical limitations that constrain spectrum usage and sharing. The first part of this book addresses the theory of spectrum sharing against the background of regulatory context and actual spectrum usage.

Chapter 1 introduces the subject with a short overview of RF spectrum usage today and how it developed over time. This serves as background for a review of current status and trends in radio regulations in Chap. 2.

Chapter 2 provides examples of policy developments by regulatory authorities, including the FCC and the UK's Ofcom, who have shown leadership in spectrum policy.

Chapter 3 provides an abstract view of the different parameter spaces, or dimensions, in which interaction between spectrum users occur: frequency, space, time, and information.

Chapter 4 uses this to provide an analysis of various ways and means of active and passive RF spectrum sharing. This includes sharing between systems with different regulatory status, as well as sharing among similar and dissimilar systems of the same regulatory status.

Chapter 5 covers basic RF technology and discusses how its various elements affect spectrum usage.

Chapter 6 covers basic techniques for spectrum sharing, notably etiquettes and medium access protocols.

Together, these chapters provide the theoretical basis for Part II, in which the practice of RF spectrum sharing for commodity wireless technologies is analyzed in detail, using examples drawn from commodity wireless technologies such as Cellular, Wi-Fi, Bluetooth, and Zigbee.

Chapter 1
RF Spectrum, Usage and Sharing

1.1 What Is RF Spectrum

This chapter gives an introduction to the current usage of RF spectrum and how that has emerged in more than a century of experimentation and practice. Given the breadth and depth of the subject, extensive coverage is not possible, nor is it necessary, given the many publications that address this field. Instead, this chapter focuses on types of usage which are related to telecommunications such as Fixed Links and Mobile Services, but radar and scientific use are also included. The chapter closes with an introduction to the two basic categories of spectrum usage: exclusive usage based on some form of spectrum property rights and non-exclusive usage based on either licensing particular types of usage or classes of users – e.g. public mobile radio – or based on exempting spectrum from licensing – e.g. the 2.4 GHz bands used by Wi-Fi and Bluetooth. How these basic elements are used in radio regulations is described in Chap. 2, Radio Regulations and Policies.

The term RF spectrum refers to that portion of the electromagnetic continuum that is subject to some form of regulated use. At its lower end, this range includes the Ultra-Low Frequencies (a few kilohertz) used for communications that span the globe and penetrate earth's (watery) surface. Its higher end reaches into the sub-millimeter waves that correspond to frequencies of 300 GHz and more.

Current usage of the RF spectrum has been shaped by practical reasons: with lower frequencies, operating range and antenna size increases but the range of frequencies decreases, making wide channels difficult to justify. With higher frequencies, the operating range that can be obtained drops, but wider bandwidths become easy, partially because the spectrum is more plentiful. Historically, the lower frequencies were the first to be used because the technologies available at the time – at the end of the nineteenth century – made that possible. The spark gap transmitter apparatus used by Heinrich Hertz in 1887, to demonstrate the existence and propagation of electromagnetic waves, may have covered a few MHz with its spectrum. Marconi's pioneering radio telegraph and the first systems transmitting audio

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_1,

signals[1] also operated within that range. With the emergence of the tube amplifier (Lee De Forest), higher frequencies became accessible and the size of radio devices and antennas could shrink to the point where a few square millimeters are sufficient for a high precision GPS receiver.

Although the highest frequencies that are covered by national or international frequency managers extend up to 2,400 GHz, the most sought after RF spectrum is a range that spans 10 octaves between ~30 MHz and ~30 GHz. However, interest in higher frequencies is rising: 60 GHz has enjoyed some popularity for high data rate point to point solutions[2] and more recently for wireless LAN applications. The IEEE 802.11 Working Group is developing a standard for 1 Gbit + wireless LANs in the frequency range of 60 GHz. However, the demands placed on the most useful lower frequencies is not likely to change because of the physical facts mentioned above and, therefore, our discussions will be limited to this frequency range.

1.2 Uses of RF Spectrum

1.2.1 Overview

The types of use made of RF spectrum span a wide range: at the lower end, long-range communications and radio/TV broadcasting predominate; at the higher end, the use of high-speed data transmission and location determination predominate.

Figure 1.1 below shows the well-known United States' Frequency Allocation Chart[3] that gives a high level view of the allocations[4] of various frequency ranges to different types of "service" in the parlance of spectrum managers. Note that this chart has a non-linear format: vertically it is logarithmic, every block represents an order of ten and, therefore, the predominance of broadcasting – suggested by the size of large blue blocks, is illusory.

There are different types of spectrum allocations: primary allocations and secondary allocations. These allocations are made to "services" – e.g. the mobile service or the radiolocation[5] service. The primary allocation affords the incumbent service protection against interference from a service with secondary status. Different rules for deciding compatibility between services in the same or adjacent spectrum apply to these allocations. In addition, regulators may allow some system

[1] See Fessenden [130].

[2] In the US, the rules for 70–90 GHz range allows very high rate point to point systems with an FDD arrangement, but little use is made of this possibility.

[3] Published by the NTIA, US Department of Commerce, see http://www.ntia.doc.gov/osmhome/allochrt.pdf.

[4] Spectrum is said to be allocated to a "service," but it is assigned to specific users or applications.

[5] Generally referred to as "radar."

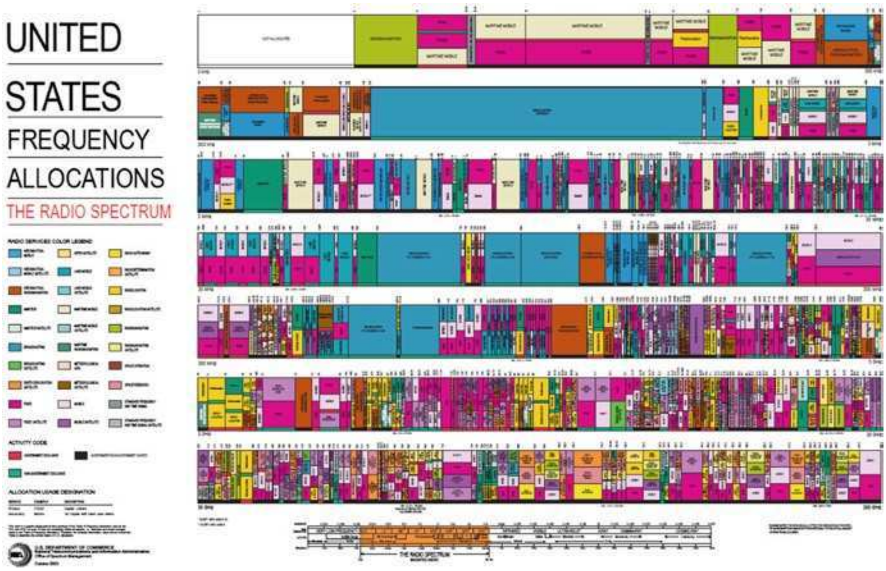

Fig. 1.1 United States' frequency allocation chart

or type of use to operate in allocated spectrum on a non-interference/non-protection basis. A service without a primary allocation is subject to interference from a service with a primary status, but not vice versa. Thus, a primary allocation is the Holy Grail of spectrum allocation for commercial service providers, as well as for safety of life services – to mention just two examples.

The non-interference/non-protection status typically applies to commodity systems – wireless LANs and Bluetooth being examples. Although much RF spectrum is not heavily used, little is unused or not assigned to some purpose and therefore shared use of RF spectrum is the rule rather than the exception. Each new use has to be judged against the interests of all concerned, including the incumbents (i.e. current users) and society at large. In some cases, the type of spectrum use requires exclusivity and thus excludes sharing. Examples are broadcasting, cellular telephony, and safety of life services. Notably, cellular telephony and its multi-media progeny have seen an astonishing rise in popularity – at the time of writing, it serves more than two thirds of the world's population.

In some cases, spectrum of sharing is possible – provided the operating parameters of the systems involved are amenable to sharing. An example is the case of radar systems and communication systems in the same frequency band. Finally, there are types of use that do not require specific parameters beyond a certain RF power level to assure acceptable sharing of the same frequencies by multiple users. A typical example is wireless LAN technology, which has seen a meteoric rise in popularity since the early years of the twenty-first century.

There are many other spectrum uses that are important and that have become part of the general infrastructure of a developed society: amateurs, broadcasting, civilian

Table 1.1 Detail of United States' frequency allocation chart

Frequency, MHz	Allocation/usage			
2,360–2,385	Mobile		Radiolocation	Fixed
2,385–2,390	Mobile			Fixed
2,390–2,400	Amateur			
2,400–2,417	Amateur			
2,417–2,450	Radiolocation			Amateur
2,450–2483.5	Fixed		Mobile	Radiolocation
2483.5–2,500	Radiodetermination (Space-Earth)		Mobile satellite (Space-Earth)	
2,500–2,655	Broadcast satellite	Mobile	Fixed satellite (S-E)	Fixed

radar, military radar and C3 systems,[6] meteorological observation, navigation systems,[7] radio astronomy and communications infrastructure using point to point links and satellite links. The 3–30 GHz range is largely divided up between the Fixed Links (both terrestrial and satellite based), the Mobile Service which is an umbrella for a wide variety of services, including Private Mobile Radio and cellular, and radionavigation and radiolocation. The latter comprise a wide variety of commercial, government and military radars systems.

Table 1.1 above shows a small part – 2,360 – to 2,655 MHz of the United States' Frequency Allocation Chart that demonstrates both the variety and density of the frequency allocations. For example, the bands 2483.5–2,655 MHz are allocated to five different types of use: fixed links, mobile, satellite-based mobile, satellite-based Earth Exploration, and satellite-based broadcasting. The license exempt wireless LANs (known in the US as "Part 15 devices") do not appear in this chart.

The 4th through 7th rows indicate the 2.4 GHz ISM band – used for a variety of applications as well as by amateurs. This band is designated for non-communication purposes – like heating with microwave radiation – and that imposes a condition on wireless services that use this band: the amateurs, radio location, the Fixed Service, the Mobile Service, etc., they cannot claim protection from interference caused by ISM applications.

1.2.2 Fixed Services

The term "Fixed Services" refers to services and systems in which both ends of a transmission link are stationary. Fixed Services provide a large part of the world's communications infrastructure, notably in cases where cable is too expensive relative to the amount of traffic carried. A good example of the importance of fixed

[6] Communications, Command and Control.

[7] Including conventional ranging systems and satellite-based systems such as GPS, Glonass, Galileo.

services are the long-range fixed links that in many countries span large tracts of countryside with high-speed digital links that carry anything from television broadcast programming to telephony and internet traffic. Over the last 20 years, these links have been joined by satellite-based links and by fiber optic cable. The latter has absorbed a large part of the new capacities brought on-line by network operators, notably in the developed economies. The more important variety of the Fixed Links is the Line of Sight Fixed Link; thanks to careful planning with respect to terrain and man-made obstacles and thanks to advanced transmitter and receiver technologies, these workhorses deliver up to 155 Mb/s over large distances at very low outage levels. Fixed Link technology is able to deliver spectrum efficiencies of up 8 bits/Hz over distances of up to 50 km, using frequencies between 900 MHz and 40 GHz. Some of this is assigned for private use and other bands assigned for use by network operators. Some examples of Fixed Link frequencies are:

2.5 GHz: licensed, MMDS (Multichannel Multipoint Distribution System)
3.7 GHz: unlicensed, Fixed Satellite Service, space to earth
5.6 GHz: unlicensed, Fixed Satellite Service, earth to space
23 GHz: licensed, commonly used for private microwave systems
28 GHz: licensed, LMDS (Local Multipoint Distribution Service)
38/39 GHz licensed, general purpose communications services, e.g. backhaul for cellular networks.

1.2.3 Mobile Services

The term "Mobile Services" refers to communications services involving at least one mobile device or entity. The ITU-R's[8] Terminology Database[9] distinguishes between four major types of mobile services: aeronautical mobile, land mobile, maritime mobile, and general mobile services such as cellular communication services.

Mobile use of radio spectrum has a long history that goes back to the early days of radio use: the benefits of the freedom of location offered by small radio systems were recognized early on. Such applications do not need exclusive use spectrum – provided that only one group of users has access to that spectrum. Examples are police and fire brigade communications and the dispatching radio system used in taxis before the advent cellular services. These systems are known variously as Professional Mobile Radio or (Private) Land Mobile Radio, Special Mobile Radio, etc.

Since the use of this spectrum is typically cooperative, there is no need for individual licenses; instead, licenses are awarded to a group or organization for a given number of channels. Similar arrangements exist for civil aviation and military communications. Both types of use cover large segments of the radio spectrum.

[8]The International Telecommunications Union, Radio section. The ITU is a body of the United Nations.

[9]See http://www.itu.int/ITU-R.

The frequencies that are used in the above mobile services typically range from a few MHz to the lower parts of the UHF band, e.g. 800 MHz. These frequencies have better propagation characteristics than the higher frequency bands.

Mobile cellular communications were first conceived by Bell Labs in 1947, but it took until 1979 for the first public system to go on-line in Japan, followed by the Nordic Mobile Telephony network in Scandinavia. Since then, cellular technology has evolved from analogue systems of the first generation (1G) that supported only voice services to the very high capacity digital systems of the fourth generation (4G) that support voice, video, and data concurrently. The frequencies used by these systems today include the 450 MHz band, the 800 and 900 MHz band, the 1,800 MHz band, and the 1,900 MHz. The 3G and 4G systems use higher frequencies in the 2,100 and 2,600 MHz bands; and further in the future, these frequencies may be extended into the 3.4 GHz range, as part of the IMT-Advanced programme of the ITU-R. The lower signal propagation distances associated with higher frequencies are an advantage to high-capacity cellular systems: the shorter range results in smaller cells and, therefore, more capacity per user.

1.2.4 Radionavigation and Radiolocation[10]

Although communications applications were the early drivers of the use of radio spectrum, navigation, notably in aircraft, was also one of the early beneficiaries of radio technology: Low Frequency Radio was used in the 1930s to support aircraft flying "on instruments." These systems emit signals that are coded according to their direction of emission. This allows a receiver to determine direction and/or location at any time. Ground-based radionavigation technology developed into a wide variety of systems that operate on different frequencies – with the longer-range systems at the lower frequencies and the shorter-range systems at the higher frequencies. Examples of two extremes are the VHF Omni-directional Range (VOR) system that operates in the VHF range and the Microwave Landing System (MLS) that operates in the lower 5 GHz range.

With that advent of satellites, truly global positioning became feasible. The default system – the Global Positioning System (GPS) of the US is – since 1996 – a dual use system that serves civilian and military purposes. It transmits on 1,772 MHz in the user segment and achieves a very high accuracy and reliability – enough to replace the MLS and other navigation tools. The European Galileo system and the Russian Glonass system have been launched to supplement – and if necessary replace – GPS.

Radiolocation – radar in everyday parlance – emerged in the middle of the twentieth century and rapidly became a dominant factor in radio spectrum usage. Radars provide visibility at long distances and under conditions that prohibit visual observation. They are used for navigation (both in fixed and mobile roles), for surveillance, for scientific purposes and for a wide variety of military applications.

---

[10] These are the formal ITU-R terms for applications of radar technology.

Whereas GPS relies on sensitive receivers that pick up the weak satellite signals from any direction, most radars rely on receiving the weak reflections from pulses transmitted with very high energy content at low duty cycles.[11] A 250 kW (=54 dBW) pulse power is typical for long-range radars employed for navigation or meteorological observation. Thanks to antenna gains in the order of 45 dB, such radars can receive the weak reflections of their own signals from objects as far as 500 km away. For a typical radar, the average power projected in a given direction is in the order of 65 dBm, which is 68 dB less than the peak power.[12] This large difference between peak and average power is a major factor in spectrum sharing: some systems and applications – like frame switched data transmission – can easily accommodate powerful short-term interference; others like TV broadcasting or Fixed Link Data Transmission cannot.

The frequencies used by civilian radars include the 1 GHz (L-)band for long-range radars, the 2–4 GHz (S-)band for terminal air traffic control, airport weather radars, and marine radars, and the 4–8 GHz (C-)band used for general weather radars, medium resolution ground surveillance radar, etc. Sharing between C-band radars and license exempt devices is discussed in Chap. 9, Spectrum Sharing with Primary Users. Military systems also use these frequencies – as well as much lower and higher frequencies – e.g. for communication with submerged submarines and for tactical air defense systems. Recent developments include vehicle radars and ground penetration radars that operate at millimeter or sub-millimeter wavelengths.[13] Ultra-Wide-Band radars are an important element in recent spectrum sharing development. UWB is addressed in Chap. 8, Sect. 8.3: Wireless LANs and Ultra-Wide-Band.

1.2.5 Scientific Uses

Another important use of the radio spectrum is scientific space research – either in active mode or in passive mode. The active model includes mapping the earth itself and oceanic wave heights with satellite-based sensing systems. These range from simple altimeters to highly sophisticated Synthetic Aperture Radars (SARs). SAR based interferometry has reached a level of precision that is able to show movements in earth's crust of a few centimeters – adequate for detailed mapping of crustal movements in earthquake-prone zones.

Passive use of the radio spectrum for space research purposes began with development of radio telescopes in the late 1940s. From these early beginnings, radio astronomy developed into a field that is able to map objects in space with great

---

[11] Some radar systems – called CW radars – use continuous transmission and operate at similar average power levels.

[12] For a 250 kW radar operating at 10 RPM, 300 pps and 2 μs pulse width.

[13] The literature on radar is vast. Wikipedia gives a good starting point under "radar" (http://en.wikipedia.org/wiki/Radar).

precision thanks to the use of Very Long Baseline Interferometry.[14] Cosmological research applications came with the measurement of the cosmic background radiation by Penzias and Wilson of Bell Labs.[15] From a spectrum sharing point of view, the importance of radio astronomy and other scientific uses of the radio spectrum lies in the fact that the very high gain of these systems make them very susceptible to man-made interference.

1.3 Spectrum Sharing

Use of the radio spectrum has caused enormous changes in the ways people and organizations communicate, conduct their business and entertain themselves. Because of the apparently limitless propagation of radio waves, early spectrum "decisions" focused on separating different types of usages and different groups of users from each other in the spectrum domain: separate frequency bands were assigned to radio telegraphy, radio broadcasting, later radio telephony, radio-location (radar), etc.

With the emergence of new types of use and the increasing importance of new applications such as the remote control, scientific observation, notably radio astronomy, telemetry, etc., the usable[16] spectrum rapidly filled up. Given the demands for ever more bandwidth by a growing number of applications, it is understandable that some think there is a spectrum scarcity that needs extreme measures to address. However, practice is different. Although there is significant asymmetry between the government controlled spectrum and "civilian" spectrum, there is no scarcity as such. There is, however, a strong desire on the part of many players in this field to crowd together in the most attractive spectrum range – between 30 and 3,000 MHz. In this frequency range, propagation is generally good and the achievable bandwidth adequate for most purposes. Therefore, usage of this spectrum has to be shared in some way. Some, like Calabrese and Benker,[17] argue that advancing technology – exemplified by Cognitive Radio – will make it possible to share all of this spectrum without prior arrangements.

Not only vested interests but also theory and practice suggest that these attractive visions are not realizable. The problem of spectrum sharing is first and foremost a technical problem, the solutions of which are bounded by the Shannon-Hartley theorem of effective channel throughput[18] and the fact that the environment, even in the desert or at sea, confronts us with propagation conditions that change with location,

[14] An introduction and pointers to more literature can be found at in Wikipedia under the heading "radio astronomy" (http://en.wikipedia.org/wiki/Radio_astronomy).

[15] In 1947.

[16] "Usable" refers to the applicable state of technology: as technology evolved, higher portions of the radio frequency spectrum became accessible at reasonable cost.

[17] See Benker [23].

[18] See Chap. 5, Sect. 5.2.1 *Transmission Signal Generation*.

time of day, the weather, etc. Common elements in all spectrum sharing situations are the fact that RF energy propagates in all directions and that the energy propagated is subject to exponential decay. For a point source radiating into free space, the decay exponent is 2 – this reflects that the RF signal propagates as a sphere of increasing size. In practice, the decay exponent can vary from less than two to more than ten – the higher values reflecting refraction, and other forms of distortion as well as shadowing and blocking that affect the propagation of the RF waves. This subject is treated in more detail in Chap. 5, The Physics of Spectrum Sharing. Suffice it here to note that the two factors above make it impossible to accurately determine which piece of spectrum is used by whom at a given location. More precisely, the presence of a transmitter – if detected at all – says little about the presence of a receiver of its signals in the same area. This inherent lack of predictability in spectrum sharing conditions is compounded by the fact that systems and applications differ in their tolerance for signal degradation. Therefore, differentiation of spectrum sharing constraints and requirements by application or type of usage may well prove to be inescapable for the foreseeable future – and beyond. Much of this book is concerned with the ways and means of spectrum sharing, so as to clarify when and where opportunistic spectrum access is feasible and when restrictive arrangements are required or beneficial.

Today, nearly all frequencies in today's "usable range" have been allocated[19] for some service or assigned to a particular application or user through a licensing process. The licensing process reflects the relative priority of a service or application. Primary allocations provide a service with protection from interference from other services, with a secondary allocation or a no-protection, no-interference allocation. Exclusive licenses tend to be associated with primary allocations, non-exclusive licenses may be associated with a primary or a secondary allocation. License exempt use typically reflects a no-protection, no- interference allocation.

1.3.1 Exclusive Use of Spectrum: Spectrum Property Rights

Since the mid twentieth century, an increasingly fierce debate has addressed the methods of allocation of spectrum. In 1943, the US Supreme Court ruled that spectrum was scarce and required careful control by the government because "there is a fixed natural limitation upon the number of stations that can operate without interfering with one another."[20]

The government proceeded with its established policy of assigning spectrum on the basis of "Uncle Sam knows best" which application or service should be provided with how much spectrum. The cost of possible mis-assignment was not recognized

[19] "Allocation of spectrum" refers to its designation for use by a defined service – like the Mobile Service – on a primary basis, a secondary basis or on a no-protection, no-interference basis.

[20] See Hazlett [64].

and therefore never analyzed. However, in 1959, Ronald Coase[21] in a groundbreaking paper, argued that spectrum property rights, acquired through some pricing mechanism would do much better than application licenses, provided those rights where non-restrictive with regard to the use of the spectrum and provided that the holders of such rights had freedom to sell and acquire such rights from other rights holders. History has proven that Coase was right, although it took until the emergence of the second generation of mobile technology in the 1990s before the FCC started auctioning spectrum licenses. Thus, the concept and practice of spectrum rights became well-established, and their advantage over application specific licensing has been proven. In other words, licensing is here to stay, possibly in the form of spectrum usage rights.

The spectrum policy pendulum has not stopped moving since then. Recent advances in radio technology and the success of wireless LANs operating in unlicensed spectrum have caused another wave of change – that is how proponents of such change see their quest for more free spectrum. The article cited above describes the above processes and its recent continuation very well. Two objections have been raised against spectrum property rights: (a) spectrum property rights would be difficult to define, and (b) such rights are no longer necessary in the digital era. Objection (a) reflects a misunderstanding of spectrum property rights – which define the rights of parties to take certain actions relative other parties. Objection (b) is based on the mistaken notion that digital technology will eliminate all interference problems. In fact, some argue that spectrum property rights should be largely abandoned in favor of spectrum commons.

This approach has been called a spectrum commons approach, because it regards bandwidth as a common resource that is subject to sharing protocols, and not a controlled resource that is always under the control of some authority, be it a government agency, a spectrum licensee, or both. However, in practice, license exempt bands are neither "open access" nor "spectrum commons," but government allocations, with rules imposed by regulators to mitigate conflicts. No radio may be used in such spectrum unless authorized by a regulatory authority, e.g. the FCC. Such authorizations typically prescribe power limits and technology restrictions.

Although "unlicensed spectrum" may carry much popular appeal and generate a fair amount of business in terms of equipment sales and services, the world-wide cellular services together generate well over $140B a year in revenue. This far exceeds the revenue generated by license exempt equipment and related products and services. Therefore, spectrum property rights will be with us for a long time to come. Changes, if they come at all, will most likely tend towards more flexibility for the rights holder than towards reducing the exclusivity of their rights or their options in using, selling or leasing spectrum. On the other hand, license exempt spectrum has a vital role to play in a technically advanced society. This is clearly evidenced by the emergence and success of Wi-Fi technology. The key to understanding this regulatory dilemma is the relationship between exclusivity and RF power allowance – and

[21] See Coase [40].

therefore in the spatial coverage – imposed by regulatory rules. High power requires a license; license exempt spectrum means power limitations.

1.3.2 Non-exclusive Use of Spectrum

Whereas cellular services and broadcasting need exclusive spectrum rights, many communication types can share the use of a given piece of spectrum. Depending on the application or service, licensing on the basis of a particular service or device may be necessary.

1.3.2.1 Shared Use of Licensed Spectrum

Shared use of radio spectrum has a long history that goes back to the early days of radio use: the benefits of mobile radio systems were recognized early on. Such applications do not need exclusive use spectrum – provided that only one group of users has access to that spectrum. Examples are police and fire brigade communications and the dispatching radio system used in taxis before the advent cellular services. These systems are known as Professional Mobile Radio or (Private) Land mobile Radio, Special Mobile Radio, etc. Since the use of the spectrum is typically cooperative, there is no need for individual licenses; instead, licenses are awarded to a group or organization for a given number of channels. Similar arrangements exist for civil aviation, military communications, etc. Both types of use cover large segments of the radio spectrum.

1.3.2.2 Shared Use of License Exempt Spectrum

License exempt spectrum is always used on a shared basis: no license being required, its use is open to anyone with the applicable equipment. License exempt spectrum is allocated on the basis of device properties, notably frequency of operation and RF power level. In some cases, there are additional requirements such as avoiding interference into other systems by passive or active means. A case in point is "Dynamic Frequency Selection" – DFS[22] for short — required in the 5 GHz band for wireless LANs.[23] Recent interest in license exempt spectrum, fuelled by the successes of Bluetooth and the Wi-Fi industry, has brought up the notion of a "spectrum commons" in which anyone could freely operate. Innovation would be stimulated, users would be able to pursue their own interests and all would be fine – thanks to the

[22] See Chap. 9.

[23] Known as U-NII devices in the US. U-NII stands for Unlicensed National Information Infrastructure.

absence of any rules or restrictions. Those opposed to such a scheme quickly identified the threat of the "tragedy of the commons": spectrum overload caused by too many users pursuing their own interest without much regard for the interests of others. The "tragedy of the commons" concept has broad applicability[24] also in the case of license exempt spectrum. However, such spectrum is subject to certain rules, and those rules change the potential for tragedy significantly. This insight has led to the idea that "commons" may have to come in different flavors, designed to support certain applications. The New Zealand Regulatory Authority formulated this as follows:

> Managed spectrum parks[25] are intended to allow access to a number of users in a common band of spectrum on a shared, and as far as possible, self-managed basis. The objective of managed spectrum parks is to encourage the efficient use of spectrum, innovation and flexibility, and provide for low cost compliance and administration.

Understanding if and how flavors of commons would address the purported tragedy threat requires understanding the various ways of spectrum sharing in more detail.

1.4 Conclusions

Given the properties of the radio spectrum, some frequencies are more desirable than others. Therefore, many types of use and many users target the same set of frequency bands – between 30 MHz and 30 GHz, and shared use of these bands is unavoidable. As discussed in the other chapters of this book, spectrum sharing is a complicated subject that involves natural laws and technical limitations, which put an upper limit on the efficacy of spectrum sharing tools and techniques. Early recognition of the need to protect spectrum users from each other led to the current radio regulations – which are being challenged by many with the argument that technology makes sharing of RF spectrum easy.

Regardless of how regulators will deal with rules for shared spectrum – whether application oriented or open to any usage — the very existence of such spectrum fuels innovation in radio technology and wireless applications. There are a myriad ways to use this spectrum, and all of these ways have to take into account that nearby there will always be other users with other equipment, operating according to different rules. That perspective of spectrum sharing is the basis for this book. Its content focuses on RF and protocol aspects of real systems and their interactions. Understanding these aspects is the key to developing not only new systems designed for spectrum sharing, but also to the development of advanced concepts like the use of game theory in the context of Cognitive Radio.[26]

[24] See Hardin [63].

[25] One "park" is at 2,575–2,620 MHz. See also Chap. 2, Sect. 2.3.4, *New Zealand*.

[26] See Berlemann and Mangold [24].

Chapter 2
Radio Regulations and Policies

Radio Regulations are a major factor in shaping the use of RF spectrum and the development of wireless technologies. Radio regulations have a long history that goes back to the early days of radio and they will remain an important factor for the foreseeable future – and beyond. Radio regulation relies on a number of tools, including licenses for the use of a given frequency band. Notwithstanding the promises of cognitive radio technology, regulation — including licensing — of the use of the RF spectrum remains necessary. Thanks to the success of commodity technologies like Wi-Fi, the perspective of radio regulatory policy making is changing. This chapter describes such changes, using the recent policies of FCC, the UK's Ofcom, the European Union, and New Zealand's Ministry of Economics as examples.

2.1 Introduction

The need for regulation of spectrum utilization was recognized early on when, in the late nineteenth century, radio communications became practice. Much of the early use was for ship-to-shore telegraphy and other long-distance telegraphy – and both need plenty of RF power. The state of the art at the time did not include sharp-edged filters and, therefore, a single transmitter could easily render a large tract of spectrum unusable for others. Therefore, the benefits of coordination became recognized quickly. Given the large distances that radio signals travel, such coordination had to be international to be effective.

The first International Radiotelegraph Conference held in 1906 in Berlin signed the International Radiotelegraph Convention, the first regulations governing wireless telegraphy. These regulations, which have since been expanded and revised by numerous radio conferences, established the principle of compulsory intercommunication between vessels at sea and the land. The output of this Conference contained, in addition to the political agreements in the form of a Convention, the first regulations governing wireless telegraphy. From these beginnings, radio regulation – now

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_2, © Springer Science+Business Media B.V. 2011

known as the Radio Regulations – grew with the rapidly evolving "radio" technology and its applications.

Today's RF technologies include a wealth of advanced features undreamt of in those early days. These features, which include sharp-edged filters, efficient transmitters, and highly selective receivers, have facilitated the use of ever higher frequencies, offering ever growing capacity to carry information wirelessly. However, demand for wireless capacity has grown apace and so the situation has hardly changed, certainly not in the lower 6 GHz of the RF spectrum in which RF signals travel long distances and retain the ability to interfere with receivers of other users. Thus, radio regulations remain as necessary as ever – even as their content may change with changing RF technology.

Even though RF signals cross national boundaries, the final responsibility for radio regulations lies with national governments. No international laws have been promulgated that supersede or override the authority of a government to determine how and by whom and what for the spectrum in its geography is used.[1] No government is willing to give up authority over a resource that has very high economic as well as military value.

The way national radio regulations are managed by Administrations varies considerably. In some countries, the civilian government is in control of all spectrum, and the military and other government organizations have to negotiate with the civil authorities to maintain or get access to "their" spectrum. The UK has gone so far as to put a price on all government ("crown owned") spectrum, so that it can be traded – with civil parties. Ofcom, the UK's Regulatory Authority, was given these powers under the Wireless Telegraphy [sic] Act of 1998. How these prices are established is not fully transparent, but two components are obvious: the operational need or importance and the cost of moving "crown operators" to other spectrum. In China, the other extreme prevails?: the military and the civilian authorities are each in control of parts of the spectrum, and changes in "ownership" require a long, politically difficult process. Other countries, like the US and France, have structure that falls somewhere between these two extremes: although jurisdictions are largely separate, internal processes are in place that allowed for the definition and execution of "joint use" arrangements, under which the military allow the use of some spectrum based on access rules and equipment requirements that offer adequate protection against harmful interference. The "DFS"[2] arrangements in the US and in France are examples.

If the preceding sounds less than straightforward, the transposition of such structures and processes to the international scene adds another layer of complexity. The International Radiotelegraph Conference of 1906 has morphed into the International Telecommunications Union – Radiocommunication Sector (ITU-R) operating under the aegis – and funding – of the United Nations. The following sections look into more detail of the organizations and policies that shape the playing field of radio

[1]This is true even for the European Union. Here the European Commission has the authority to issue "EC Decisions" that the member states are required to implement – unless they have overriding national considerations; in practice, such "derogations" are only temporary.

[2]See Chap. 9, Sect. 9.1, *Sharing the 5 GHz band – Radar systems and wireless LANs.*

technology developers and their customers: operators, service providers, businesses, and consumers. Since it is not possible to describe each country in detail, a few examples will serve.

2.2 International Coordination: The ITU-R

The ITU Radiocommunication Sector (ITU-R)[3] is one of the three sectors of the International Telecommunication Union (ITU). The ITU[4] is a large agency of the United Nations that addresses global issues of information and telecommunications technologies and that serves as a focal point for governments and industry in developing networks and services. In addition, the ITU, through its ITU-R sector, has promoted international cooperation regarding the allocation of radio spectrum. Since the final authority over the radio spectrum lies with national governments, the role of the ITU-R is advisory rather than law giving. Its Radio Regulations assist in the international management of radio-frequency spectrum and satellite orbit resources "in order to avoid harmful interference between radio stations of different countries." The international spectrum management system is based on regulatory procedures for frequency coordination, notification, and registration.

Every three or four years, the ITU-R organizes the World Radio Conference (WRC), at which the Radio Regulations are reviewed and revised as necessary and new ITU-R Recommendations are formally approved. Thus, the international "law" is kept approximately in line with the practice of the day in the national jurisdictions or their regional equivalent – e.g. the EU. The ITU-R Recommendations and supporting Technical Reports are prepared in the "Study Periods" that precede each WRC. These are managed by the Radiocommunication Assemblies through the assignment of preparatory work (ITU-R Questions) to ITU-R Study Groups. The Radiocommunication Assemblies also approve and issue the ITU-R Recommendations that form the technical backbone of the Radio Regulations.

Given the inherently global nature of radio communications, the function of the ITU-R is indispensable in avoiding conflict between nations about radio spectrum usage – not only in border areas. Also, the rapid globalization of product development and sales makes it necessary that radio regulations are global in nature. This also extends to the newest trends and technologies in this area such as spectrum commons[5] and Cognitive Radio.[6] In adopting these issues in its work programme, the ITU-R facilitates the global distribution of the newest developments in spectrum management.

---

[3] See http://www.itu.int/net/about/itu-r.aspx.

[4] See http://www.itu.int/net/about/index.aspx.

[5] See example e.g. Horvitz [67].

[6] Per Resolution 956 [COM6/18] (WRC 07): Regulatory measures and their relevance to enable the introduction of software defined radio and cognitive radio systems.

2.3 National Regulatory Policies – Some Examples

In general, there is a widespread trend towards lightweight spectrum management, reducing the role of the regulator in spectrum management. Some refer to this as the demise of the "command & control model" of spectrum management; others, more realistically, see this as a policy adjustment that fits with the general trend towards "less government, more market." It has the advantage of reducing the liability of the spectrum regulator for mismatches between demand and supply. In the following sections, the policies and actions with regard to license exempt or lightly licensed spectrum of a few major regulatory authorities are presented.

2.3.1 The FCC

The Federal Communications Commission of the US government was established by the Communications Act of 1934. It was preceded by the Federal Radio Commission and it absorbed wire communications responsibilities to become the key factor in the setting and execution of the US telecommunications policies. Its responsibilities include the coordination of communications policies in the Americas through its International Bureau. From a wireless communications perspective, the Wireless Telecommunications Bureau is important as well as the Office of Engineering and Technology, which supports the decision making of the Wireless Bureau with technical expertise. It also supports the Enforcement Bureau through its Equipment Authorization Branch, which oversees the regulatory compliance for all equipment operating in spectrum between 9 KHz and 300 GHz. Other Bureaus – e.g. the Public Safety and Homeland Security Bureau, have little or no impact on matters related to spectrum sharing.

The specific policy objectives for the FCC, over the period 2006–2011, are according to its 2006 Strategic Plan[7]:

> ***Broadband****: "All Americans should have affordable access to robust and reliable broadband products and services. Regulatory policies must promote technological neutrality competition, investment, and innovation to ensure that broadband service providers have sufficient incentives to develop and offer such products and services."*
>
> ***Competition****: "Competition in the provision of communication services, both domestically and overseas, supports the Nation's economy. The competitive framework for communications services should foster innovation and offer consumers reliable, meaningful choice in affordable services."*
>
> ***Spectrum****: "Efficient and effective use of non-federal spectrum domestically and internationally promotes the growth and rapid development of innovative and efficient communication technologies and services."*

[7] See for the full text http://www.fcc.gov/omd/strategicplan/.

Media: *"The Nation's media regulations must promote competition and diversity and facilitate the transition to digital modes of delivery."*

Public Safety and Homeland Security: *"Communications during emergencies and crisis must be available for public safety, health, defense, and emergency personnel, as well as all consumers in need. The Nation's critical communications infrastructure must be reliable, interoperable, redundant, and rapidly restorable."*

Modernize the FCC: *"The Commission shall strive to be highly productive, adaptive, and innovative organization that maximizes the benefits to stakeholders, staff, and management from effective systems, processes, resources, and organizational culture."*

Although these objectives mention defense, the real control over much "government spectrum" is exercised by Departments and other Agencies, e.g. the Department of Defense, the Federal Aviation Commission, etc. Coordination among these fora is facilitated by the National Telecommunications and Information Administration. Where government and commercial spectrum usage overlap or prove adjacent, the NTIA and the FCC will together develop policies. An important resource in this context is the NTIA's Office of Spectrum Management, the responsibilities of which include "providing the technical engineering expertise needed to perform specific spectrum resources assessments and automated computer capabilities needed to carry out these investigations.[8]"

2.3.1.1 The National Broadband Plan

The policies developed and executed by the FCC extend far beyond the scope of this book; and therefore, only the immediately relevant to spectrum sharing are mentioned here. First of all, there is the National Broadband Plan of 2010. It was introduced in the press[9] as follows:

In 10 years, 90% of Americans will have affordable access to 100 Mbps broadband, with schools, hospitals and army bases getting 1 Gbps access, if Congress adopts the national broadband plan published today by the Federal Communications Commission (FCC). To make this happen, the FCC will free up and make available 500 MHz of radio spectrum, remove barriers to entry for existing and new broadband providers, and improve the amount and quality of information available on broadband markets. "The National Broadband Plan is a twenty-first century roadmap to spur economic growth and investment, create jobs, educate our children, protect our citizens, and engage in our democracy," said FCC chairman Julius Genachowski.

This ambitious plan is a continuation of FCC policies over the previous decennia.

[8]For more information see http://www.ntia.doc.gov/osmhome/osmhome.html.

[9]Ian Grant [34].

2.3.1.2 The Incubation of Wireless LANs

In 1981, the Federal Communications Commission, under the direction of Michael Marcus of the FCC staff, started exploring ways to permit more general civil uses of spread spectrum. The FCC's proposals were generally opposed by spectrum users and radio equipment manufacturers. Nonetheless, the May 1985 decision[10] permitted unlicensed use of spread spectrum radios in three bands at powers up to 1 W. The rules were codified as 47 CFR 15.247 and they became a model for similar regulations in other countries and regions. These "Part 15 rules," as they became known, created a fertile breeding ground for consumer-oriented wireless technologies like Wi-Fi, Bluetooth and others, including cordless telephones. Wireless LANs became a poster[11] child of government fostered innovation. The breakthrough came when Apple Computer Inc. offered "Wi-Fi" for the magic price of $99. Public confidence was fostered by the efforts of the Wi-Fi Alliance to assure the interoperability of wireless LANs across many vendors. Wireless LAN technology generated enormous benefits for consumers, businesses, and institutions – thanks to the confluence of two other technologies: standardized computer platforms and cellular phones.

Intermezzo: Success Does Not Need a License

Motivated – in part – by the success of wireless LAN, some[12] have argued that all RF spectrum should be free and not burdened by rules or fees. The arguments put forward by the proponents of "free spectrum" largely hinge on equating "rules" with "financial burdens." Therefore, free spectrum is assumed to automatically lead to more innovation, better spectrum usage and more value, at least for some. Although the emergence of the wireless LAN technology seems to support these assumptions, reality was and is different. The rapid rise of wireless LAN technology was enabled by innovation that leveraged the relative freedom offered by the FCC's simple rules for devices operating in the 900 and 2,400 MHz ISM[13] bands. Although necessary, that freedom was not a sufficient condition – the latter was provided by the confluence of two commoditization trends: the commoditization of laptop computers by the computer industry and the rapid rise of mobile telephony. The convenience of the cell phone created the demand for wireless computing – a demand that could be met thanks to the extensive use of standard interfaces and communication protocols in the computer industry.

On the other hand, the rise of cellular telephony and its evolution into mobile multi-media services is an equally impressive example of how licensed RF spectrum can form the basis of a vast new industry – one that has caused a pervasive

[10] See http://www.marcus-spectrum.com/documents/81413RO.txt.

[11] Used here in the sense of archetype.

[12] E.g. the New America Foundation.

[13] ISM = Industrial, Scientific and Medical [applications].

change in the way people communicate. It stands to reason that the assurance of clean spectrum for a long period of time allowed the huge investments needed to create the large cellular networks and the highly sophisticated technology for handsets, smartphones and, not to forget, the backhaul transmission systems. This point is most clearly demonstrated by the GSM system developed in Europe in the late 1980s. Following a Memorandum of Understanding between 13 operators and administrations that proposed the development of a common cellular system, the EU provided an allocation of spectrum to facilitate roaming throughout the continent. In 1993, the GSM system had a million subscribers served by 70 carriers in 48 countries. As of the second quarter of 2009, GSM and its derivates served 3,450,410,548 connections. Its major competitor, CDMA-2000, served 528,696,917 users.[14] In 2007, the global revenue of cellular operators exceeded $140B.

The following sections introduce four policy initiatives of the FCC that illustrate the willingness to adopt unconventional approaches to the management of the nation's spectrum resources and stimulating innovation. These initiatives are Ultra-Wide-Band, Interference Temperature, lightly licensed broadband in the C-band and Cognitive Radio.

2.3.1.3 Non Plus Ultra – The Saga of Ultra-Wide-Band

Another example of policies that aim to stimulate innovation is the Ultra Wide Band ruling of the FCC. The core concept of UWB is that by spreading energy in frequency rather than in time, the RF power seen by existing systems would be so low as to be below the background of emissions by other – intentional and non-intentional[15] – transmitters. Since channel capacity increases linearly with bandwidth – for the same SNR – the throughput of UWB was promised to exceed any wireless technology known at the time.

On September 21, 1998, the FCC issued a Notice Of Inquiry[16] titled "Revision of the Rules Regarding Ultra Wide Band Systems" (UWB), which proposed, along with unlicensed radar systems that are able to see through walls and into the ground, the unlicensed use in the range of 3.1–10.6 GHz by UWB transmitters for communications purposes, provided that the power spectral density of the emissions would be below −41.3 dBm/MHz.

Some hailed UWB as basis for a new revolution in unlicensed communications and one that could rival or exceed the success of wireless LANs; others thought the proposals to be a serious threat to currently deployed systems and services because ubiquitous deployment of UWB devices would raise the RF noise floor sufficiently

[14]Figures for 2009, reported by the GSMA through their GSMWorld website.

[15]These terms refer to true RF transmitters designed to radiate energy and to devices that produce radiation as "by product" of their operation – e.g. stray energy from electric motors and leakage from microwave ovens.

[16]See FCC 98-208, 1998, Revision of the Rules Regarding Ultra-Wideband Transmission Systems.

to cause loss of operating range. The ensuing debates and controversies raged for years, and the IEEE 802.15 Working Group that was set up to develop a standard for UWB failed to meet its objectives. However, the FCC persisted and in its February 14, 2002 First Report and Order, the FCC authorized the unlicensed use of Ultra Wide Band systems for data transmissions.[17]

The R&O stated that "UWB technology holds great promise for a vast array of new applications that we believe will provide significant benefits for public safety, businesses and consumers. With appropriate technical standards, UWB devices can operate using spectrum occupied by existing radio services without causing interference, thereby permitting scarce spectrum resources to be used more efficiently."

The technical requirements were based on studies and recommendations of the NTIA. The R&O document also noted that "[t]his has been an unusually controversial proceeding involving a variety of UWB advocates and opponents. These parties have been unable to agree on the emission levels necessary to protect Government-operated, safety-of-life and commercial radio systems from harmful interference."

In 2005, the first devices were demonstrated, and in 2007 the first products were certified by the FCC. Today, in 2010, UWB is the basis of Wireless USB – a technology promoted by the WiMedia consortium.[18]

2.3.1.4 Interference Temperature

At the time it issued the UWB Report & Order, the FCC also announced its Notice of Inquiry on the concept of "interference temperature" as a measure of spectrum utilization and as a tool of spectrum management. It had been proposed in the report of an internal staff study – under the heading of "Spectrum Policy Task Force," commissioned in 2002 and promised a revolution in spectrum management practice. The NOI[19] introduced the matter as follows:

> *This new concept could shift the current method for assessing interference which is based on transmitter operations, to an approach that is based on the actual radiofrequency (RF) environment, taking into account the interactions between transmitters and receivers. The interference temperature model could represent a fundamental paradigm shift in the Commission's approach to spectrum management by specifying a potentially more accurate measure of interference that takes into account the cumulative effects of all undesired RF energy, i.e. energy that may result in interference from both transmitters and noise sources, that is present at a receiver at any instant of time.*
>
> *This new approach could provide radio service licensees with greater certainty regarding the maximum permissible interference, and greater protections against harmful interference that could be present in the frequency bands in which they operate. In addition, to the extent that the interference temperature limit in a band is not reached, there could be*

[17] The same R&O also provide rules for wall and ground penetrating radar applications and automotive radar applications of UWB technology at higher frequencies – see FCC02-48A1.

[18] See http://www.wimedia.org/.

[19] See FCC 03-289A1.

opportunities for other transmitters, whether licensed or unlicensed, to operate in the band at higher power levels than are currently authorized. In such cases, the interference temperature limit for the band would serve as an upper bound or "cap" on the potential RF energy that could be introduced into the band.

The ensuing discussion was full and vigorous. There was some spirited defense of the concept[20] and there was criticism. The proceeding was closed without further ado in May 2007.[21] The main reason for its demise was that the proponents ignored the fact that measuring the average interference at one point says little about that level at another point. This is due to the ever changing, unpredictable environment on signal propagation. Although not a success by any means, this proceeding does illustrate that the FCC was willing to go to great lengths to meet its stated goal of enabling innovation in wireless technologies through spectrum management measures.

2.3.1.5 Lightly Licensed Broadband Operations in the C-Band

The C-band, including its extensions, ranges from 3,400 to 6,425 MHz. It is much used by satellite-based broadcasting and communications services. The downlink frequencies range from 3,400 to 4,200 MHz, the uplink from 5,850 to 6,425 MHz. Notably in the industrialized world, fiber infrastructure is providing most of the capacity needed for long-distance communications and this has reduced the demand for satellite-based communications.

In 2005, the FCC released 50 MHz of spectrum at 3,650–3,700 MHz for shared use under a light licensing regime.[22] The initial considerations go back to the Report and Order of October 2000 that allocated these 50 MHz on a co-primary basis to the fixed stations and base stations of the Mobile Service and to Docket 02-380, which concerned additional spectrum for unlicensed devices below 900 MHz and in the 3 GHz band. Given the low density of use of these 50 MHz, the FCC decided to make it available for those wireless ISPs that sought to provide wide area coverage, but who were confronted with lack of suitable spectrum. The Report and Order allows fixed stations in this band to transmit with up to 25 W in as many MHz: the power density limit was thus set at 1 W/MHz. For mobile devices, the limit was set at 40 mW/MHz.

With the success of IEEE 802.11-based devices in mind, the FCC decided to require that these systems implement a "contention-based protocol" so as to avoid interference – and therefore minimize regulatory oversight. Major protests ensued – mainly from vendors and prospective users of TDMA systems. As a result, the FCC modified their rule such that e.g. WiMAX systems could be operated in this band as well; the modified rule allows "restricted contention protocols" in the

[20] Kolodzy [77].

[21] See FCC 07-78.

[22] See FCC-05-56A1.

lower 25 MHz. In this case, the term “restricted” refers to the property that the contention protocol works only between like systems such as WiMAX.[23] Chapter 9 considers other co-existence issues related to WiMAX and other 4G technologies. Following the FCC’s announcement of the new rules, the IEEE 802.11 Working Group quickly ratified an Amendment (“y”) to their Wireless LAN standard that contains the technical requirements necessary for “wireless LANs” to operate under the new rules. This opened the way for the application of CSMA/CA technology to networks with much larger internode distances, as foreseen by the initiators of the standard.

The FCC noted that there was no evidence that these “contention-based protocols” would be in adequate to avoid interference among the users of this band and, therefore, they added a registration requirement: all fixed, outdoor (base) stations operating in this band must be registered with the FCC in the Universal Licensing System data base. Such registration allows prospective users to find out what systems have been deployed already and it allows parties to find each other in case of interference. The latter point is important because the FCC stated that parties are assumed to resolve interference issues – which admittedly could arise because of the absence of exclusive access to this spectrum – among themselves. Recognizing that interference could also arise from mobile devices, the Report & Order requires that mobile devices shall not transmit unless they receive an enabling signal from a fixed – and therefore registered – base station.

Another issue associated with this band is the presence of operational ground stations of the Fixed Satellite Service. Although few and far between, they do merit protection and the Report & Order defines a number of exclusions zones around the sites to be protected. Although the above appears to present a unique opportunity for wireless service providers to deploy systems with good coverage but without a costly spectrum license, practice turns out to be more complicated. Field experience[24] shows higher than expected interference and an ineffective registration system – because the latter does not require the registrant to declare the channels used.

2.3.1.6 TV White Space Spectrum in the US

The growth of wireless services has raised the demand for more spectrum to be made available. Such spectrum is becoming available as the result of the analog to digital switch-over of TV transmitters: digital TV transmission packs more channels in the same bandwidth – thanks to different modulation techniques and higher transmitter power levels. In 2004, the FCC released a Notice of Proposed Rulemaking which identified the possibility of using some of the TV frequencies in UHF bands

[23] Note that the same applies to Wi-Fi devices: most of these do not use energy detection, but carrier detection – the former allows for some degree of spectrum sharing with other types of system, but the latter is far more precise.

[24] As niche operator WEBFORMIX found out, this is even true in semi-rural Oregon, USA.

for communications purposes. Given the way TV channels have been allocated in the past, it was not possible to create a single, contiguous band for communications applications. This "patchwork" aspect gave this spectrum its informal name: TV White Space spectrum. Similar policies were pursued in Europe and elsewhere.

Spectrum at these low frequencies has been used for TV transmissions for good reason; for broadcasting it has highly desirable properties: less attenuation with distance, less blockage by obstacles and better wall penetration. Those same properties are less advantageous for telecommunications use: the lower propagation loss means that not only the operating distance increases, but also the interference caused by other networks. Therefore, the capacity of this spectrum – in bits/second/area – is much lower than the capacity of spectrum at higher frequencies. This consideration favors services that require coverage rather than capacity.

In the US, the freed-up television frequencies are primarily in the upper UHF "700-MHz" band, covering TV Channels 31–69 (698–806 MHz). In total, 105 MHz would become available. In the rest of the world, some of the abandoned television channels are being reallocated for Digital Audio Broadcasting in its various forms. In the US, there was considerable opposition of the broadcasting community and wireless microphone users to having to share the VHF and UHF frequencies with unlicensed devices. In November 2008, the FCC released a Second Report & Order on the subject that ruled in favor of license exempt devices sharing with the incumbents. However, the rules also imposed stringent measures on the former to protect the latter, based on a combination of database look-up and sensing. Subsequently, the sensing requirement was removed,[25] although the FCC note that:

> *[…] While the Commission is eliminating the sensing requirement for TVBDs, it is encouraging continued development of this capability because it believes that it holds promise to further improvements in spectrum efficiency in the TV spectrum in the future and will be a vital tool for providing opportunistic access to other spectrum bands.*

For the technical aspects of TV white space spectrum sharing see Chap. 9, Sect. 9.3.2 Sharing TV White Space Spectrum.

2.3.1.7 Cognitive Radio

The final policy example to be mentioned is the FCC's Cognitive Radio proceeding. It was heralded by some as the beginning of the end of spectrum scarcity.[26] The concept was first put forward by Mitola in 2000[27] and blessed by the FCC's Spectrum Policy Task Force Report of 2002.[28] The proceeding started with the issuance of a Notice of Inquiry – FCC 03-322 – in December 2003 in Docket 03-108. The NOI

[25] See the FCC 08-260A1, Second Memorandum, Opinion and Order.

[26] See Staple [117].

[27] See Mitola [96] (http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-970).

[28] See ET Docket 02-135.

stated in its introduction that cognitive radio technology would improve spectrum usage in a variety of ways and situations, including automatic frequency coordination among licensees of co-primary services, by facilitating coordination within a licensee's network, by facilitating coordination between licensees and third parties in secondary spectrum markets and by allowing license exempt use of licensed spectrum at times or locations it is not or lightly used.

In a way, this can be read as a retrenchment from the earlier abstract and less mature "Interference Temperature" concept towards a more practical approach. The Cognitive Radio model allows flexibility to address the problem of spectrum sharing, while avoiding intervention by the spectrum manager or regulator. However, the fundamental role of the FCC as the spectrum assignment authority was not changed at all. Given the loud criticisms – from some quarters – of the FCC's licensing policies, this may be considered somewhat of a surprise.

2.3.1.8 Summary

The preceding sections show that the FCC, in addition to its traditional role of spectrum regulator through licenses of various kinds, has a pro-active policy of establishing new spectrum usage criteria for license exempt spectrum. These criteria are based on new concepts, and they are intended to leverage new technologies and procedures such as contention-based protocols, registration of spectrum users and, notably, Cognitive Radio. Whether this policy will deliver the expected benefits remains to be seen, but the challenge for the wireless industry is clear.

2.3.2 UK Regulator Ofcom

The UK's Ofcom or "Office of Communications,"[29] as it is known formally, is an agency of Her Majesty's Government. It was established by the Communications Act of 2003. It inherited the responsibilities of a number of other bodies, including the Broadcasting Standards Commission, the Office of Telecommunications, and the Radiocommunications Agency – the latter being its parent with regard to radio regulations and technical standards. With regard to radio regulations, its main areas of activity include research, licensing, policy development, addressing complaints and looking into competition.

Ofcom advertises its role as follows[30]:

> *One of Ofcom's key statutory duties is to ensure the optimal use of the radio spectrum under its management. Radio spectrum is a major asset to the UK, contributing some £24bn to the economy each year and underlying many aspects of our lives. Radio communications is critical to areas such as air travel, emergency services, cellular telephony, sound and television broadcasting, defense and our utilities.*

---

[29] See http://www.ofcom.org.uk/.

[30] See the Introduction at page http://www.ofcom.org.uk/consult/condocs/sfr/sfr2/.

Ofcom actively develops spectrum usage policies and its preferred vehicle is the public consultation. This process is somewhat akin to the FCC's Notice of Proposed Rulemaking: a Consultation typically announces a policy or policy objective and presents Ofcom's own thinking on the subject, together with the intended rulemaking and a set of questions that seek public comment on its objectives or rulemaking. The responses obtained are then used to form the final decisions – which are usually published with the responses received. In addition to its national role, Ofcom has an international role in representing UK's interest in international fora, like the ECC and the ITU-R.

2.3.2.1 Ofcom's Main Spectrum Policy Initiatives

One of the first actions of the newly constituted Ofcom organization was to develop and publish its "Spectrum Framework Review" (SFR). In the SFR, which is heavily influenced by the findings of the Cave Report,[31] it sets out the background and future of its radio spectrum policies.[32] The framework identifies three ways to manage the radio spectrum: the conventional "command and control" style management, which is used for 95% of the spectrum, market-based spectrum management within boundaries set by Ofcom, and license exempt use based on RF power limits and other technical constraints as necessary. The latter method is used for some 5% of the spectrum. The SFR serves as a vehicle for Ofcom to change the balance between these different approaches. The guiding principles given in the SFR consultation can be summarized as:

The SFR identified the following means to achieve its goals:

- Spectrum for license-exempt use is needed, and will grow to seven percent of the total spectrum.
- The implementation of trading and liberalization, where possible, will allow market forces to prevail. This policy will be applied to 72% of the spectrum.
- The remaining 21% of the spectrum will be managed using current practice of licensing.

The above goals reflect a transition from a technically oriented style of spectrum management to an economics-oriented style. The interests of citizens and consumers are assumed to be best served by competition among providers of equipment and services. This approach is solidly embedded in the Communications Act of 2003. An open, non-discriminatory approach and transparent licensing rules are fundamental aspects of the Act, which aims to harmonize and simplify license rules and conditions wherever possible. This approach has given the UK a leading role in Europe regarding secondary trading, implementation of digital TV, introduction of "light touch" licensing (such as lifetime licensing service for ships' and amateur radio), with a technologically neutral format for spectrum auctions to facilitate liberalization.

[31] See http://www.spectrumaudit.org.uk/final.htm.

[32] See http://www.ofcom.org.uk/consult/condocs/sfr/sfr2/.

2.3.2.2 Ofcom's License Exempt Framework Review (LEFR)

The LEFR[33] Consultation closed in June 21, 2007 after the usual 10-week review period. In the Consultation, Ofcom pointed out that the contribution to the UK economy of license-exempt applications is significant. As an example, they assess that the net present value of public Wi-Fi local area networks (without taking congestion and interference costs into account) might be as high as £100 bn over 20 years. By some estimates, this is only a quarter of the net present value that could be generated by licensed cellular networks over a similar period. Nonetheless, this figure emphasizes the importance of license-exempt use of the radio spectrum, and the need for an appropriate form of management.

The LEFR Consultation sought answers to a number of questions, including some which are relevant to the purposes of this book. Ofcom published its findings, noting, among others, that:

(a) *"Where possible, any spectrum released in the future for license-exempt devices should be used based on the spectrum commons model, wherein devices for multiple applications share the same frequencies, subject to politeness rules and polite protocols. While spectrum commons will be the default approach, exclusive license-exempt use of spectrum by a specific application will be considered on a case-by-case basis where technical constraints, international obligations, or safety issues require such use."*

(b) *"Multiple classes of spectrum commons should be considered, with regulatory-defined mandatory politeness rules restricting the diversity of applications within the bandwidth of each class (thereby easing co-existence), and with polite protocols micro-managing intra- and inter-application interference. The specification of polite protocols should be undertaken by appropriate technology standardization bodies or individual equipment manufacturers, and is beyond the scope of the regulatory bodies."*

(c) *"Light-licensing regimes should only be adopted when explicit co-ordination among the operators of the radio devices is both feasible and a technical necessity (i.e. when limitations in technology prevent autonomous self-coordination among the devices). Licence-exemption should be adopted otherwise, subject to adequate protection of incumbent users."*

(d) *"Radio devices transmitting at sufficiently low power spectral densities do not cause harmful interference to incumbent services, and should be exempted from licensing. A power spectral density lower bound for the licensing of radio devices should be considered which: (a) is equal to the UWB limits for frequencies below 10.6 GHz; and (b) is extrapolated from the UWB limits for frequencies above 10.6 GHz (accounting for increased signal attenuation with frequency). Transmissions below the specified limits may be exempt from licensing[..], subject to compliance with all UWB operational restrictions (other than minimum bandwidth) as specified in EC Decision 2007/131/EC."*

(e) *"[..] Harmonised technical standards are expected to be sufficient for mitigating the impact of interference caused by compliant radio transmitters, particularly at high frequencies where radio propagation conditions and the abundance of bandwidth imply a low probability of congestion."*

The LEFR Consultation was followed by Studies and Consultations on politeness rules, spectrum commons rules and EIRP limits for license exempt devices. These are summarized in Appendix A.

[33] http://www.ofcom.org.uk/consult/condocs/lefr/summary/.

2.3.2.3 Ofcom's Policy on Spectrum Usage Rights

The major piece of regulatory re-engineering of Ofcom is the shift away from licensing of applications and specific systems towards tradable spectrum licenses that give the licensee a wide degree of freedom in the use and/or trading of the licensed spectrum. Although neither spectrum trading nor licensing policies are subjects of this book, the management of interference at the edges of the licensed spectrum of geographical space contains elements of spectrum sharing that are of interest.

In order to achieve the necessary flexibility in use and trading, a new way of license specification was needed. The existing licensing policy typically used the technology, the application or mode of use as parameters. Examples are 3G mobile licenses and PMR licenses. Wanting to leave the licensees – and their sub-licensees – as much freedom as possible while protecting the interests of users in adjacent bands or geographical areas, Ofcom focused on restrictions at the boundaries addressed by the license. These restrictions are defined in a new instrument called "Spectrum Usage Rights" or SURs.[34] Licenses based on SURs are specific to a given frequency band and apply to a given geographical area. This area can be as large as the UK.

The SURs are defined by three components:

(a) the aggregate power flux density at a given boundary should not exceed a certain level in dBW/m^2/MHz below a given height above local terrain for more than some percentage of the time;
(b) the out-of-band power flux density should not exceed a certain level in dBW/m^2/MHz below a given height above local terrain for more than some percentage of time and at more than some percentage of locations in a given area;
(c) the in-band power flux density should not exceed a certain level in dBW/m^2/MHz below a given height above local terrain for more than some percentage of time and at more than some percentage of locations in a given area.

Component (a) assures separation from geographical neighbors operating in the same frequency band, whereas the other two components assure separation in the spectrum domain.

Both simple and elegant, the above SUR definition is useful for transmitters serving certain areas, e.g. base stations and fixed links. Put simply, only the downlink to adjacent downlink interference case is considered. However, mobile point-to-multipoint systems consist of multiple transmitters and may use TDD or FDD protocols. Ignoring this additional complexity carries a price in efficient spectrum use (because wide guard bands would be needed). In an update of 2009,[35] Ofcom addressed this and proposed to add the following modeling cases to the downlink-downlink case already covered:

- uplink to an adjacent uplink
- downlink to an adjacent uplink
- uplink to an adjacent downlink

[34] See http://www.ofcom.org.uk/consult/condocs/sur/.

[35] See http://www.ofcom.org.uk/consult/condocs/surs/.

In May 2008, Ofcom added a "License Verification"[36] approach that spells out how to verify compliance to the SURs obtained by a licensee. Such compliance verification would be performed to answer a complaint from another licensee. Ofcom proposed to run simulations based on real transmitter locations and accepted propagation models to determine the expected power flux densities at the license's boundaries.[37] Responses to this Consultation raised many technical objections and it remains to be seen when and how Ofcom proceeds with the transition from traditional licenses to SUR-based licenses.

2.3.2.4 Conclusion

The preceding sections sketch out the audacious policy changes that Ofcom is pursuing. Most notable are the Spectrum Usage Rights model of licensing and the spectrum commons classes approach, for allocating license exempt spectrum based on the notion of a (supposedly neutral) Interference Indicator.[38] Both promise significant liberalization of spectrum usage that many have asked for, but, notably in the case of the spectrum commons, at the price of reduced certainty. Provided adequate attention is paid to the physics of RF transmission and propagation in implementing these policies, the desired results may well prove achievable. However, given the complexity of spectrum sharing analysis – as exemplified by this book – caution is advisable.

2.3.3 The European Commission

This short section gives a quick overview of two key elements of the EU's spectrum policy. This policy, although general in nature, works in the same direction as those of the US and the UK.

2.3.3.1 Introduction

The European Union has identified radio spectrum policy as a tool to realize – or help realize – the ambitious goals of the "Lisbon Agenda" set in 2008.[39] The responsibility for spectrum-related policies lies with the directorate "Information Society." The thematic portal[40] of the Information Society directorate describes this policy as follows:

---

[36] See http://www.ofcom.org.uk/consult/condocs/surs/statement/.

[37] The EC defines Block Edge Masks for the purpose of band sharing. For an example see http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:144:0077:0081:EN:PDF.

[38] See Appendix A.2.

[39] See http://europa.eu/scadplus/glossary/lisbon_strategy_en.htm.

[40] See http://ec.europa.eu/information_society/policy/ecomm/radio_spectrum/eu_policy/index_en.htm.

At the European level, radio spectrum policy has three main objectives:

- *Harmonizing the use of radio spectrum*
- *Working towards more efficient use of the spectrum, and*
- *Improving the availability of information about the current use of spectrum, future plans for use and availability of spectrum.*

A primary aim of EU policy in this area is to ease access to spectrum usage by reducing bureaucracy and delegating decision-making to the users where possible. To achieve its aims, the Commission has outlined a forward-looking strategy for Radio Spectrum policy in Europe. The strategy advocates a common EU-wide approach, using a mix of models to manage frequency allocation as appropriate to the type of wireless application envisaged.

The approach includes increasing the supply of spectrum by ensuring, through a market-based approach, that assignment of spectrum to individual use is efficient. This approach is often optimal in commercial sectors where Quality of Service (QoS) is important, such as mobile telephony.

At the same time, flexibility is increased in some bands by making their use unlicensed, i.e. allowing the assignment of individual rights to be withheld for applications or for parts of the spectrum where such rights are not justified and shared usage represents the best use of resources. The deregulated use of spectrum fosters innovation in rapidly evolving mass-markets such as short-range devices (e.g. WiFi) where a degree of harmful interference is acceptable. Traditional management processes, where the use of the spectrum is tightly regulated by governments, will also continue to be applied selectively in areas where security and safety-of-life considerations are paramount (e.g. defence, public security, aviation). Two general principles are applied in EU spectrum policy: technology neutrality and service neutrality. In essence, all innovations, products, and processes are deemed to be equal.

2.3.3.2 General Wireless Access Policy – WAPECS

One tool in the EC policy tool box the WAPECS approach: Wireless Access Policy for Electronic Communications [and] Services. According to the Commission, the deployment of innovative wireless services and technologies is hampered by the reservation of certain useful frequency bands for a narrow set of services. In combination with rigid usage conditions, this is assumed to delay the application of new technologies and the efficient use of [the available] radio spectrum.

Under the WAPECS approach,

avoiding interference will remain a key element of spectrum management, but the way this can be achieved has evolved due to rapid technological progress. The main objectives of WAPECS are to achieve improved technical and economic efficiencies in the market. To prepare for this new approach, a step-by-step introduction of flexible spectrum management is required. This entails identifying specific spectrum bands where regulatory restrictions can be lifted to introduce competition: including competition between different radio infrastructures. A total of 1,350 MHz of bandwidth has been identified as a first set of bands in which legal restrictions could be re-examined to permit more flexible usage. The bands are currently used by a variety of broadcasting, mobile and information technology sectors. An EU-wide set of proportionate rights and authorization conditions must be agreed that would represent the minimum necessary to allow flexible and efficient usage without interference between services. These conditions would also need to facilitate the gradual readjustment of existing rights (legacy rights) held by spectrum users.

The above 1350MHz of spectrum includes:

- *470-862 MHz: used for broadcasting today;*
- *880-915 MHz / 925-960 MHz as well as 1710-1785 MHz / 1805-1880 MHz: these bands form the 900/1800 network for GSM mobile services today;*
- *1900-1980 MHz / 2010-2025 MHz / 2110-2170 MHz; these bands are used for third generation (3G) mobile services (IMT-2000/UMTS) today;*
- *2500-2690 MHz (the 2.6 GHz band); this band (still to be licensed) is intended for use by 3G mobile services (IMT-2000/UMTS);*
- *3.4-3.8 GHz: this band is used for broadband connections to customers; it is also intensively used for satellite communications in Russia and some African countries."*

Interestingly, the WAPECS approach assumes that "competition between different radio infrastructures" will lead to more efficient spectrum use. Much of this book is concerned with the physical and technical aspects of shared spectrum use. Its results point in another direction: spectral efficiency is best served by homogeneous use of spectrum access methods and protocols. This means that competition between "radio infrastructures" is unlikely to achieve the Commission's goals. The reason is that the first entrant into some piece of spectrum is likely to make that spectrum unattractive to other candidates.

2.3.3.3 Digital Dividend Spectrum in Europe

The WAPECS programme includes the TV White Space Spectrum, also known as the Digital Dividend (spectrum). The UK led the European developments and has nearly completed the digital switchover which freed up spectrum of the channels 31–40 and 62–69 – a total of up to 128 MHz, depending on the release of channels 39 and 610. Channel 38 will remain blocked to protect radioastronomy services. Studies by Ofcom[41] and the CEPT[42] indicated that the economic benefit of license exempt spectrum use in these bands was exceeded by the economic value of licensed use. This example corresponds to the gist of a report[43] developed for the European Commission on the subject; it also found support in many other countries. Thus, there is little prospect for license exempt use of the Digital Dividend spectrum in Europe and the 47 member states of the CEPT. Instead, primary usage will be digital terrestrial television and "broadband wireless services," e.g. LTE and WiMAX, both allowed under the stringent neutrality policy of the EU with regard to technology and services. Whether these systems will use cognitive technologies is another matter. An ECC report is available[44] that addresses the use of cognitive technologies for opportunistic sharing of the 470–862 MHz range. Its conclusions are close to those of the FCC:

[41] See http://stakeholders.ofcom.org.uk/binaries/consultations/ddr/report_analysys1.pdf.

[42] See ECC Report 159 [4].

[43] See http://www.analysysmason.com/PageFiles/13825/20090814%20EC%20final%20-report%20-%20Final%20(Executive%20summary%20only).pdf.

[44] See ECC Report 159 [4].

sensing is not good enough because of the technological limitations,[45] database look-up is a better means of coordinating spectrum usage.

2.3.3.4 The Collective Spectrum Usage Model

Another policy tool of the Commission is the "Collective Use of Spectrum" model. This model allows shared use of spectrum and is assumed to lower possible barriers to the use of radio spectrum.

CUS allows a number of spectrum users to share spectrum in through separation in frequency, space, and/or time. Although associated with license exempt usage, the CUS model "should be seen in a broader context [...] 'Light licensing' regimes (registration or notification), underlay (via ultra-wide band (UWB) technology) and overlay[46] (cognitive radio) approaches and possible future types of spectrum sharing (for example via 'private commons') are also captured under this approach. Most Short Range Devices (SRDs) also use spectrum in accordance with this model. [...] Under the CUS approach there are collective usage rights instead of individual (exclusive) usage rights and interference management is dealt with by spectrum regulation (containing technical parameters) in combination with self-regulation between users. This replaces interference management by a spectrum management authority through spectrum regulation and license conditions under the licensed model. The CUS approach allows for a reduction of technological and regulatory constraints and increases the responsibilities of spectrum users to share spectrum efficiently, manage interference effectively and even to accept limited interference in certain cases."

As described here, the CUS approach includes license exemption and dedicated commons as regulatory tools. The latter can be dedicated to a technology, a type of application or usage or a set of users. So far, the CUS approach has not been applied to specific frequencies.

2.3.3.5 Conclusion

Clearly, the European Commission looks at radio spectrum policy as a tool to foster growth in the Union. Its policies follow the same track as those of the US and the UK, although in a more generic sense. This approach, while enabling new developments, avoids proceeding on the basis of assumptions that may prove not to hold up in practice.

---

[45] Actually, physical impossibility is the more important reason.

[46] Some use the terms *underlay network* and *overlay network* in the context of spectrum sharing. We avoid such use of these terms because their proper use refers to network architecture. These terms are used when one wants to distinguish between the physically determined aspects – the underlay network – and the logically determined aspects – the overlay network – of a communications systems.

Table 2.1 License types used in New Zealand

Type of license	Main property	Example
Assigned license	Exclusive, time limited license with precise geographical and frequency boundaries	2G/3G license for mobile operators
Non-assigned license	Non-exclusive, power-limited license for a frequency range	PMR
General user license	Non-exclusive license with power and frequency limits	Wireless LANs, SRDs, etc

2.3.4 New Zealand

New Zealand may lie far from nearly every country in the world – including Australia – but its spectrum management policies, as described in the "Report Review of the Radio Spectrum Policy in New Zealand,"[47] line up well with the advanced policy discussed above. Like the UK's, its Radiocommunications Act and Telecommunication Act recognize different ways or managing access to non-government radio spectrum; these are shown in Table 2.1.

The Assigned Licenses need not be tied to a single licensee: licenses can be traded in the secondary spectrum market. The other two types of license will be impacted by the technical developments that facilitate spectrum sharing. The Report notes that Cognitive Radio does not fit very well in the above scheme.

In 2008, New Zealand's Ministry of Economic Development embarked on a unique project known as the Managed Spectrum Park or MSP[48] for the 2.5 GHz band. The underlying concept is to allow users – in this case operators – to work out among themselves the best way of using the frequencies available in the MSP. The main characteristics of the MSP regime are:

(a) The MSP [license] is available only for local and regional services covering a maximum of ten contiguous local areas, such as Auckland City, Auckland Region, etc.
(b) Users determine the specifications of licenses — this to assure technology neutrality and flexibility.
(c) Licenses have a six-year term with a right of renewal on certain conditions; use-or-lose provisions apply.
(d) A "resource rental" is charged for use of the MSP.

[47] For a good overview see http://www.rsm.govt.nz/cms/policy-and-planning/completed-projects/competition.

[48] See http://www.rsm.govt.nz/cms/pdf-library/policy-and-planning/radio-spectrum/managed-spectrum-parks/msp-discussion-paper-kb-pdf, also available as ISBN 978-0-478-31652-0.

Where the demand for licenses in an area is such that quality of service is potentially compromised, a period is set for applicants to coordinate and agree on revised specifications. If they cannot do so, the Ministry will draw lots to progressively eliminate applications, until the remainder can be accommodated.

Applicants are required to declare:

- Proposed service and technology,
- Characteristic frequency,
- Frequency band (lower and upper bands),
- Maximum power (in relation to horizontal radiation pattern – degrees), Location of transmitter (for a fixed base or central hub of a point-to-multipoint or mobile service, or for each part of a point to point service, this must be specified as a location, not an area),
- Site name and height,
- Antenna polarization,
- Antenna height.

Experts will evaluate these parameters and advise on the technical feasibility. In case there is more demand than supply, all applications have to share their data, propose modifications and submit to arbitration. This also applies to new entrants entering an MSP after initial users have been established.

This experiment points to another method of allocation of spectrum in which cooperation between license holders, incumbent and prospective, is the basis for the licensing of spectrum to regional interests. Experience will tell whether the MSP is successful, but the need for understanding spectrum sharing issues and their careful evaluation is clear.

2.4 Looking Forward

2.4.1 Licensed Spectrum and Spectrum Usage Rights

Spectrum licensing is here to stay, even as the regimes under which licenses are issued change to allow more influence of the parties involved. This clearly comes across in publications used as background for the preceding sections. Also clear is the trend towards licensed spectrum trading. Licenses – with different conditions and operating rules – will become tradable items and the rights attached to them will become more flexible with time. The UK's policy of making more than 70% of the non-government spectrum available under tradable Spectrum Usage Rights is unique in the world and it remains to be seen if it will prove successful. Success in this case is linked to trust in the Spectrum Usage Rights concept and as well as to proof of their efficacy in practice. In principle, there is no reason to doubt the value of Spectrum Usage Rights as a spectrum management tool because the domain of application is mostly static in frequency and time: once boundary conditions have been determined to the satisfaction of the parties involved, there is little prospect of unexpected change – the bane of most commercial operations.

2.4.2 Commons – Non Specific License Exempt Spectrum

License exempt spectrum will remain the smaller part of the spectrum allocations in most countries, regardless of how licenses are handled. The value of license exempt spectrum has been proven as breeding ground for new technologies, but – according to economic theory, the economic value remains below that of licensed spectrum. The rules under which such spectrum is made available will vary by country, but it may well be that the UK's idea of spectrum common classes finds a following, possibly in another format or with different class criteria. The advantage over "non-specific commons" is that a class-based commons avoids putting incompatible systems – referred to as "dissimilar systems" in this book – together in the same spectrum. As this book makes clear, the classification scheme that underlies the class-based commons scheme should be expected to fall short of its goals. Absent large-scale practice, its effectiveness – or lack thereof – remains a matter of conjecture.

2.4.3 Cognitive Radio and Dynamic Spectrum Access

Dynamic spectrum access – the Holy Grail of the Cognitive Radio community – is certainly on the agendas of many regulatory authorities,[49] as well as on the agendas of researchers and entrepreneurs. Its promise is to enable the use of spectrum that is underused – either geographically or time-wise – through Dynamic Spectrum Access (DSA) and its cousin Opportunistic Spectrum Access.

Because of its regulatory implications, Cognitive Radio has made it to the agenda of the ITU-R, where it is being addressed in Study Group 1. The initial results suggest that Cognitive Radio techniques will not develop into a "service" or application that deserves its own spectrum or regulatory rulemaking. Instead, the ITU-R has concluded that the application of Cognitive Radio techniques must be considered on a case-by-case basis.

Some see Cognitive Radio techniques as enablers of spectrum commons – an idea that has already seen partial realization in the protocols and spectrum access protocols of license exempt technologies. The DSA toolbox of Cognitive Radio contains a variety of elements, including frequency agility, spectrum sensing, location awareness, spectrum usage databases and spectrum usage rules, transmit power control and – interestingly – negotiated spectrum usage. Notably for the military, DSA offers great potential, in that it can help expeditionary systems to rapidly adapt to local conditions and opportunities. Cost and complexity being less of concern and spectrum sharing requirements limited to the "own kind," military DSA systems are feasible today and they are in fact being built and deployed.[50] Commercial

[49] See e.g. Anker [13].

[50] DSA technology should not be confused with Software Defined Radio technology – the latter is a technique for sharing hardware and software across a number or "radio applications" each of which operates according to conventional, typically licensed spectrum access rules.

DSA is another story – it may sometime reach maturity, but that time seems far off and the conditions required for reaching maturity are uncertain. Spectrum sensing is a key component in a DSA system, not only in operation but also the creation of spectrum usage data bases. It is a recent "invention" that may prove amenable to development to the point where it becomes a reliable tool. However, given the non-linearity of RF signal propagation and the variability of the propagation environment, spectrum sensing may well prove of limited practical use beyond certain limited cases. Nonetheless, spectrum sensing will remain a feature of DSA systems, together with other capabilities.

Much effort has been put into developing an understanding of how the benefits dynamic spectrum access could be realized – e.g. through cooperative spectrum sharing. Until the practicality of dynamic spectrum access has been proven, the advanced game theoretical designs for cooperative spectrum sharing are likely to prove valuable; but until such time, they will only be applicable to a limited set of deployment conditions, e.g. short-range indoor communication scenarios.[51]

[51] See Berlemann [24], p. 207.

Chapter 3
Dimensions of Spectrum Sharing

All spectrum sharing takes place in various dimensions or parameter spaces: frequency, information, space and time or in some combination thereof. In the following, each of these dimensions will be considered in more detail with examples drawn from established technologies and systems.

There are two modes of spectrum sharing: passive and active. In the passive mode, arrangements are static or change very slowly in time. In the active mode, devices decide dynamically on the actual arrangement, i.e. the parameter values for each of the four sharing dimensions.

3.1 The Frequency Dimension

In the early days of radio transmissions, it was clear that the long-distance propagation of radio signals was both a boon and drawback: not only the intended receivers could receive the signal, but also receivers that belonged to other systems could receive a transmitter operating on their frequency. It was quickly found that a simple way to achieve separation among different radio systems is to share a common spectrum by separating in the frequency domain – by selecting different bands or channels, they would see less interference. Thus, sharing RF spectrum in the frequency dimension was the first method of spectrum sharing to be developed.[1] Although the efficiency of channelization depends on the selectivity of the filters used in the transmitter as well as in the receiver, in general, channelization facilitates predictable and efficient spectrum sharing. Efficient channelization comes at the price of forcing all sharing systems to use the same channel width – or at least the same basic raster.[2]

---

[1] In time, this led to the allocations of spectrum to specific types of use and this developed into the complex regulations we have today.

[2] This is seen in standards like IEEE 802.11 that provide for channels to facilitate sharing with its own kind without providing means to share with other types of systems.

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_3,

In the case of passive channelization, the frequency channels are pre-assigned – each transmitter has its own channel and receivers know which channel to tune to in order to receive a certain transmitter's signals. Changes in such an arrangement require prior coordination and are slow to be implemented; and therefore, this form of spectrum sharing is typically associated with licensing that provides exclusive spectrum usage rights. The most obvious example is broadcasting. Adding separation in frequency by direction of information flow as in Frequency Division Duplex systems (FDD) improves sharing options as well as overall spectral efficiency: the same frequencies can be re-used at shorter distances. This fact is exploited in cellular mobile systems. The downside of FDD operation is that it is potentially inefficient: if the traffic load varies in direction or time, the frequency utilization varies as well. In the case of IP data traffic, such variations can be very large and the spectral efficiency suffers correspondingly.[3] However, when traffic patterns change slowly – i.e. when they are roughly constant during a call or session, FDD is an efficient way to share spectrum passively. FDD should not be confused with FDM,[4] which provides for the parallel use of one (broad) channel by multiple users.

In the case of dynamic sharing, channel selection takes place at the time of use. That selection can be based on prior knowledge or on active sensing. Prior knowledge requires that the transmitter obtains information about the locally available channels from a data base or from a beacon control channel. A case in point is the DECT system: its channels have two dimensions: frequency and time; which combination to use for a given client connection is determined by the DECT base station at the time the connection is established.

Using active sensing, a radio system could perform measurements on different frequencies prior to selecting a frequency, to determine which frequencies are free to be used. Radio devices that are capable of selecting one channel out of the set of available channels, improve overall spectrum utilization considerably. In some scenarios, the overall spectrum utilization can be improved at little or no cost. A prime example of such dynamic sharing is the operation of Wi-Fi devices in the 5 GHz band: they are required to sense the presence of radar systems before being allowed to use a given RF channel. This scheme is known as DFS.[5] Its adoption opened the way to use 455 MHz of mostly clean spectrum for license exempt applications. Chapter 9, Sect. 9.1 describes DFS in detail.

Regardless of the scheme of frequency sharing that is employed, there is always the bandwidth ratio to be taken into consideration. The simplest case is if interferer and victim have the same effective bandwidth: detection and avoidance parameters are the same and analysis is simple. There are four different combinations of interferer and victim bandwidth and these yield four different cases of sharing in the frequency domain, as shown below.

Table 3.1 points out that, if significant bandwidth overlap cannot be avoided, there is an asymmetry in interference potential between narrowband and wideband

[3]This argument is the primary justification for cellular TDD systems like WiMAX.

[4]See http://en.wikipedia.org/wiki/Frequency-division_multiplexing.

[5]Dynamic Frequency Selection, see also Chap. 9.

Table 3.1 Impact of bandwidth differences on interference

		System A (interferer)	
		Narrowband	Wideband
System B (victim)	Narrow band	Proportional relationship between detectability and interference sensitivity	Proportional relationship between detectability of A and its interference sensitivity – little impact on victim
	Wideband	Inverse relationship between detectability of A and its interference sensitivity – large impact on victim[6]	Proportional relationship between detectability and interference sensitivity

systems that cannot be overcome on a purely bi-lateral basis.[7] Given comparable power levels, narrowband systems will always "damage" wideband systems, and not the other way around. The wideband systems can avoid such damage only by trading increasing the robustness of their signaling at the expense of data throughput. That consideration lies behind the original Part 15 Rules of the FCC for devices using spread spectrum techniques.[8] In the proceedings on docket 99–231,[9] the industry successfully argued that much higher performance would be possible without wideband spreading rules – without mentioning the implications of losing the robustness afforded by these rules. The FCC agreed to remove these "unnecessary" spreading rules, motivated by two considerations: the potential increase in the data rate of wireless devices and, secondly, there was no change in interference potential relative to other users of that spectrum. The relaxed rules opened the door to dramatically increased data rates and the use of OFDM and MIMO techniques. These high data rates do little to balance the loss of robustness against interference. In other words, the performance achieved by Wi-Fi devices is realized only in the absence of pervasive interference.[10]

The frequency domain can be used to facilitate multiple access: many users sharing the same resource on an as-needed basis. This is known as FDMA — users are assigned to one or more channels for a given period, typically on the basis of actual demand. The advantage is that the usage of the channel(s) is not burdened by time or distance constraints. Conventional FDMA is not very efficient for data heavy applications – it is not able to rapidly adapt to changes in usage patterns and traffic loads. However, the combination of multiple access and OFDM has breathed new

[6]Within certain limits: if the BW ratio is large and B uses OFDM, it may be able to handle narrowband interference.

[7]However, if the probability of frequency overlap is taken into account, the asymmetry is much reduced see Chap. 5, Sect. 5.1.1. *Transmitter Signal Generation*.

[8]See Section 15.247 – 902 to 928 MHz, 2,400–2438.5 MHz, 5,725–5,850 MHz, (Spread Spectrum and Digital Modulation Specified) (Bluetooth, WiFi, Digital Cordless Phone).

[9]See also FCC 02-151 of May 16, 2002.

[10]Self interference of Wi-Fi devices is largely avoided through the CSMA/CA medium access protocol.

life into this form of spectrum sharing. It is called OFDMA and it is used in cellular systems. See Chap. 6, Sect. 6.5: *OFDMA: Multiple Access with OFDM*.

A variation of FDMA that offers more flexibility is the frequency hopping spread spectrum (FHSS) mode of operation: time is divided into slots and different users get different frequencies in each slot. As in the case of simple FDMA, steep filters are needed to keep the information flows of different users clearly separated. An added requirement is the need to synchronize the hopping pattern between transmitter and receiver. A downside of FHSS is the inherent bandwidth limitation: supporting many users or sessions requires a large number of hops and this reduces the bandwidth available per hop. Complex schemes of hop aggregation have been designed, but few have seen application in widely used products. More detail is provided in Chap. 4, Sect. 4.2. Techniques for Spectrum Sharing.

The above discussion does not include the selectivity aspect: in practice, transmitters and receivers are not perfect, but show profiles in emitted RF power and sensitivity that go well beyond the frequencies of interest. Chap. 5, Sect. 5.3.4. Receiver Selectivity discusses the implications.

3.2 The Information Dimension

Sharing RF spectrum in the information dimension exploits the information carried by radio signals for sharing purposes. Radio transmissions may carry information in many ways. The simplest case is given by an un-modulated carrier transmission – its presence or absence provides the receiver with a single bit of information. When modulation is added, more bits of information are carried, depending on the form and complexity of the modulation. Adding data increases the information content of the signal, but it does not help the receiver to recover the signal. That process is facilitated by adding redundant information[11] to the transmitted signal. By mixing the data and the redundant signal, the required bandwidth is increased, but the required signal to noise ratio is reduced. This reduction is also called processing gain; it can be exploited for multiple access purposes – as in Code Division Multiple Access (CDMA).

In CDMA, one code word (chip sequence) is shared between two stations – typically a base station and a mobile device. If the separation between the code words is sufficiently large, such pairs of stations can communicate at the same time without suffering interference from each other. The spreading rate required depends on the net capacity required and, therefore, it depends on the required SNR. For 8 kb/s voice, a margin of 17 dB (SNIR + fading) may be adequate and that means a spreading rate of at least 50. The UMTS[12] system uses variable spreading with rates ranging from 4 to 512. The higher values come into play at less favorable transmission conditions. Since the latter include the number of active users on a given channel, CDMA may not be suitable for truly high bandwidth, mobile multi-media

[11] If this redundant information is known in advance by the receiver, quick detection is facilitated.

[12] Third Generation cellular communications system.

applications. In the UMTS system, the spreading rate can be as high as 333, which corresponds to a processing gain of nearly 25 dB. The CDMA technique of spectrum sharing is most effective if all the users in a given frequency band use it. In that case, the design parameters can be tuned to achieve a desired operational profile – e.g. one that optimizes multi-user behavior at the expense of operating range.

3.3 The Space Dimension

Sharing RF spectrum in the space dimension exploits the fact that the properties of the space through which an RF signal propagates affect the properties of that signal. These spatial properties include e.g. the presence or absence of attenuating material, reflections and blocking obstructions; they determine the signal's strength and phase at any point. This fact has benefits as well as downsides. The former include the possibility of an almost infinite variety of signal propagation paths between any two points in space. That subject is addressed below in Sect. 3.3.3 MIMO: Leveraging Multiple Propagation Paths. The downsides include loss of operating range and link reliability.

3.3.1 Distance

As discussed in Chap. 5, Sect. 5.2.3 Signal Propagation, propagation plays a major role in spatial spectrum sharing: in free space signal attenuation is only 6 dB/octave – this means that every distance doubling reduces the signal strength by 6 dB. For indoor situations and in heavily built-up areas, the attenuation is in the range of 10 dB/octave or more. The resulting difference in propagation range – both from the viewpoint of operating range and interference range – is huge as shown in Fig. 3.1.

However, unless the properties of the traversed space are known in detail, it is impossible to relate the properties of the received signal to the distance that the signal has travelled. Because of this uncertainty, a large safety margin is required in estimating the distance to a given transmitter. This makes the use of the distance factor for sharing of spectrum is inefficient. In urbanized settings and within buildings, physical separation goes a long way towards spectrum sharing, notably among low-power devices. This is widely proven in practice by the success of license exempt devices – e.g. Wi-Fi and Bluetooth – operating in the crowded 2.4 GHz band: the RF power levels of these devices are so low that the signal strength drops below the noise floor at relatively short distances – a few tens of meters in most buildings.

Figure 3.1 points towards potential pitfalls of spatial spectrum sharing – whether planned or managed. Because of the exponential signal strength degradation, most of the coverage area of a given transmitter receives a signal that is close to the receiver's limits. In the case of free space propagation, the outer 3/4th[13] of the

[13] The difference in range is ½, the difference in coverage is $(½)^2$.

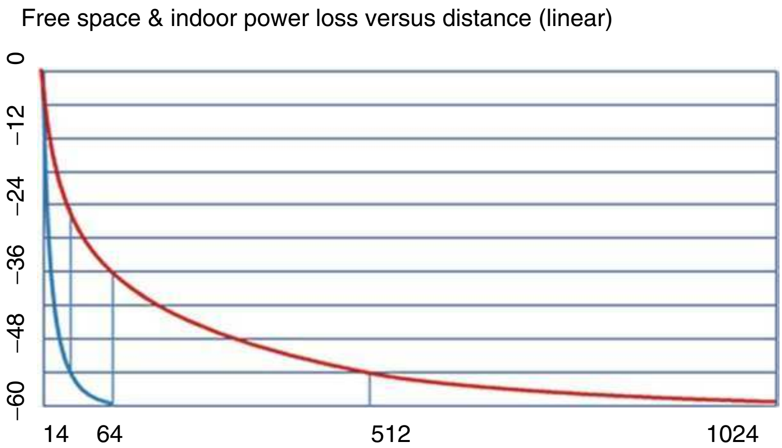

Fig. 3.1 Signal attenuation with distance

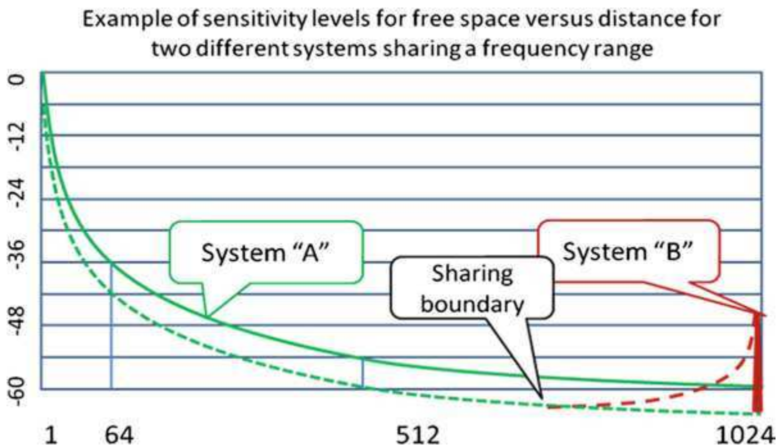

Fig. 3.2 Example of the sharing boundary between two systems

coverage area receives a signal that is only 6 dB weaker than the signal in the inner one fourth. For a higher pathloss exponent, the outer zone is even larger. Therefore, small errors in the assessment of coverage contours may result in large areas being underserved or not served at all. This is illustrated in Fig. 3.2.

In this example, the SNIR of System A is assumed to be 6 dB. If System B is located at the very edge of System A's operating range – i.e. where it could just receive System A – its signals would cause a severe loss of range to System A – some 30% in the case that System B operates at 30 dB less RF power output than System A. In order for System B to avoid any interference (= any range degradation) to System A, it has to sense System A's signals at least 6 dB below the noise floor of the System A receivers.

Because of these uncertainties, spatial sharing based on sensing is not a solution for systems with an appreciable operating range in free space, nor in case there is a large difference between the bandwidths of the systems concerned. If the interferer B is a narrowband system and the victim A is a wideband system, B has to sense A's signals at $(6+10\log\Delta BW)$dB to avoid reduction in A's operating range. However, in case of dissimilar systems sharing the same band, ΔBW may not be known and a (large) safety margin will be needed. Even if low-cost technology will make sensing at such low levels possible, the uncertainties caused by unknown bandwidth ratios and unknown propagation factors will lead to large margins of uncertainty. The result is either significant probability of interference or – in case protective margins are applied – inefficient spectrum utilization.

3.3.2 Directional Antennas

An effective way of utilizing the space dimension to share spectrum is to use directional antennas. The benefits of directional antennas are twofold: the tighter the beam they produce, the smaller the area in which interference is caused to other systems and, secondly, the reduced chance that this narrower beam will intersect the location of the receive antenna of another system. The former aspects are covered in detail in Sect. 5.2.3. The coupling probability is minimized if the potential victims also use directional antennas. In that case, the overall chance of interference is proportional to the square of the antenna directionalities: two antennas with $\alpha=3°$ will have a $(1/120)^2=69.44\ 10^{-6}$ chance[14] of seeing each other and to interfere with each other. Clearly, sharing spectrum using high gain directional antennas is very efficient.[15] However, the improved sharing behavior comes at the price of reduced spatial freedom. This can be compensated for by means of beam steering; that subject falls outside the scope of this book.

3.3.3 MIMO: Leveraging Multiple Propagation Paths

Advances in channel modeling and signal processing have opened up another ways of exploiting the spatial dimension: by making use of channel diversity to carry multiple signals at the same frequency. This gain can be used to increase the robustness of the information being carried or to reduce the time needed to transport a given amount of information.

Rich physical variety in the space between a transmitter and receiver means that there is a large number of paths between them that may be virtually independent. To use such a path, transmitter and receiver require a separate antenna. A number

[14] Expressed in dB this is −42 dB.

[15] Some Administrations specify RF power limits in terms of EIRP – this is exceedingly conservative and limits the use of license exempt devices unnecessarily.

of these paths taken together form a "multiple input/multiple output channel" that supports multiple data streams between two physical locations. There are various ways of using multiple input/output channels: they range all the way from one input to n outputs (SIMO) to n inputs and one output (MISO). The same applies to devices: one can think of multiple receivers optimized for receiving the signal from a single transmit antenna and, at the other extreme, one receiver that is optimized for receiving signals from multiple transmit antennas. The former - SIMO - type of channel configuration is a bit like conventional receiver diversity. Applying Maximum Ratio Combining to a 1×4 SIMO configuration provides considerable reduction in the SNIR required for a given PER. 1×4 MRC can provide as much a 10 dB gain relative to a 1×1 SISO system at a PER of 10^{-2} the same data rate.

The above clearly shows the increased link robustness due to the lower SIR requirement that can be obtained with MRC (a similar result applies to STBC). The main implication of a drop in the required SIR of 10 dB is that, in an indoor environment, the number of potential interferers is reduced by a factor 4. This, in turn, points towards a gain in throughput of a factor 4 – this offsets the drop in raw data rate necessary to obtain the necessary link robustness. In other words, dropping data rate and leveraging MRC at the receiver reduces the effect of interference and smoothes the link behavior, possibly to the point where it behaves like wired connection.

3.3.4 Summary

The preceding discussion clearly shows that the optimum way to make use of the spatial dimension for spectrum sharing purposes is the use of directional antennas that exploit the reduction in interference "footprint" around a transmitter. Since this applies equally to transmitter and receiver, the reduction in interference is potentially very large for systems with similar, high gain antenna systems.

The next best method is the use of MIMO technology that exploits the presence of multiple paths between two points to increase either raw throughput or signal robustness or some combination of the two. As with directional antennas, the maximum gain is obtained when all systems use similar transmitters and receivers.

The least effective way to use the spatial dimension for sharing spectrum is to rely on distance between transmitter and victims. This works well only in case of low-power devices, notably when combined with MIMO technology. The downside is compensated for by the generality of the method.

3.4 The Time Dimension

Sharing RF spectrum in the time dimension exploits the fact that many transmission systems operate in a non-contiguous manner – packet data transmission, the very basis of the Internet, is a prime example. Transmission in the time dimension means

giving up continuity in exchange for freedom from interference – at least in principle. There are many ways to do this; most come down to the same thing: avoid transmitting when that could interfere with another transmission in the neighborhood. The avoidance can be static or dynamic; the latter case splits in a distributed control variant – e.g. Wi-Fi – and a centralized control variant – e.g. WiMAX.

3.4.1 Sharing without Information Exchange

There is a passive variant of the distributed approach: the Aloha family of protocols. In Aloha protocols, collision avoidance requires that the channel utilization is limited so that collisions may be expected to be infrequent enough to be recoverable by retransmission. See Chap. 6, Sect. 6.1 Simple Etiquettes: Aloha and "Listen-Before-Talk." Another form of passive sharing would be pre-assignment of time slots that are distributed in advance to the users sharing the channel. Such a system can hardly be considered practical for use in license exempt spectrum because there is no way to share the assignment data.

3.4.2 Sharing by Explicit Information Exchange

3.4.2.1 Distributed Medium Access Control

A more flexible approach than pre-allocation is actively sharing a channel between a group of cooperating users by means of a token. The users decide how to use a channel access token: whoever has the token can send and/or pass the token on if so desired. This scheme allows adaptation of the allocation of transmission time to users and therefore it is an improvement over fixed time allocations. Token sharing on a wired bus was never a success, token sharing over wireless is wasteful; even if there is no useful traffic, there is still token traffic. The lack of reliability of wireless links requires another layer complexity to be added to this scheme, which makes it even less attractive.

Another downside is that schemes like these require a common protocol for channel access which is contrary to the philosophy which underlies the provision of license exempt spectrum. However, it could be used in an "application specific commons."

3.4.2.2 Centralized Medium Access Control

TDMA systems also require a common channel access protocol: all participants in a cell – e.g. a point to multipoint arrangement – have to use the same signaling (for channel usage related information), so that they all know what is going to happen next. This aspect limits the application of TDMA systems to spectrum that is an application specific commons.

Another aspect of TDMA systems is that they require special provisions to share the same channel among multiple, co-located or nearby cells. TDMA controllers have to know the load levels of their neighbors and execute a protocol to resolve or at least reduce conflicting channel demands. An example of a scheme designed to provide distributed spectrum resource sharing is the Spectrum Load Smoothing (SLS) developed by Berlemann and Walke.[16] SLS has been proposed as a Cognitive Radio solution to cooperative channel sharing. It requires participating stations – in this case TDMA cell controllers – to exchange load information, so that all can determine how much to decrease – or increase – their use of the common channel so as to achieve nearly constant loading over time. SLS could provide the claimed benefits, but only in conditions that ensure cooperation. This limits its application to a spectrum commons that is application specific.

The token passing approach fails on efficiency and reliability and TDMA requires a complex inter-cell coordination mechanism. Both are limited to use in application specific spectrum. Therefore, other ways are needed that do not suffer these drawbacks. The following addresses only the case of systems that do not need explicit information exchange for medium access resolution.

3.4.3 Channel Sharing Etiquettes[17]

Procedures for distributed control of channel sharing are generally known as listen-before-talk etiquettes. For such etiquettes to be effective, it is necessary that all participants keep track of what others are doing and, secondly, that all make their actions known – either implicitly or explicitly. Sharing information is easy if all devices are of the same or similar "kind": in that case they can understand each other's signaling and actions. If they are not, matters get more complicated.

The required information consists of two sets of data: the first set describes which participant[s] is[are] currently transmitting and for how long, the second set describes the expected participant[s] and its[their] transmission[s]. The first set is easily obtained in real time: the stations interested in acquiring the channel listen to the channel and collect the necessary information. The second set requires that all the acquiring stations publish their intent to acquire the channel and that some mechanism is applied to resolve the unavoidable conflict that arises because participants are "deaf" while they transmit.

The whole problem is simplified a lot if the number of transmitting participants is limited to one: its transmission informs the others about the duration of its occupancy of the channel and, at the end of that time, synchronization occurs after which only one of the others is allowed to transmit.[18] Thus, the end of the current transmission is the synchronization point for all participants.

[16] See Berlemann and Mangold [24], p. 164 etc.

[17] An etiquette does not require explicit information exchange; a protocol defines such exchanges.

[18] Note that this limitation is artificial: in practice, propagation conditions are such that there will always be stations who do obtain the necessary information and therefore transmit unexpectedly and possibly destructively.

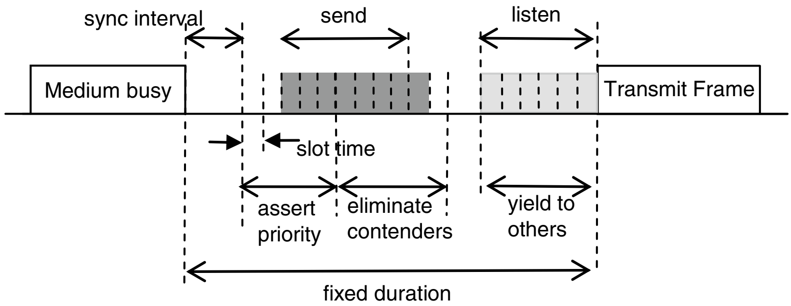

Fig. 3.3 The three phases of the EY-NPMA medium access scheme

3.4.4 Synchronization and Prioritization of Channel Access

The problem of ordering all "candidates" is reduced to establishing who wins the contention for the channel. This can be resolved in two ways: by explicit signaling and by random trial & error. The first approach was chosen by the HIPERLAN 1 standard,[19] the second by the IEEE 802.11 wireless LAN standard.[20]

The HIPERLAN1 standard, developed by ETSI in the 1990s, defines a combination of pilot tone transmission and listening to reduce the population of active contenders to one. The Elimination Yield-Non Pre-emptive Multiple Access (EY-NPMA) scheme makes use of three phases as shown in Fig. 3.3. In the Prioritization phase, the devices first to transmit silence all other devices. In the Elimination phase, stations that transmit longest, have the best chance of surviving – survival being validated by listening — the Yield phase — before the transmission phase begins. The phases are short, the probability distribution for a given "actions" is exponential. This combination makes the scheme quite stable and robust under heavy load. The signaling overhead is fixed and has to be efficient relative to the typical frame duration. By adapting notably the length of the Elimination signaling, its efficiency for short frames can be significantly increased. Since no explicit information is exchanged, the scheme can be applied to dissimilar systems: it can be used in general license exempt spectrum that is not application specific. A downside to the EY-NPMA scheme is that it relies on energy detection rather than on modulated signal detection: the latter is more precise and is, therefore, better able to defer to weak signals and so avoid being subject to interference from remote devices.

The IEEE 802.11 Distributed Coordination Function (DCF) scheme takes a different tack: stations synchronize on the end of the current transmission and begin to transmit after some random interval called the back-off window. Down counting of the interval is suspended whenever a station hears another station's transmission Fig. 3.4.

[19] Anastasi [12].

[20] Ferre et al. [46].

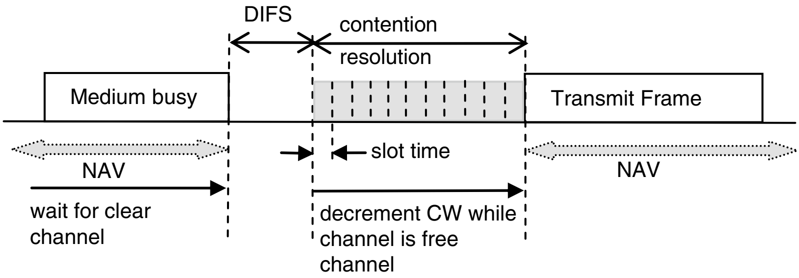

Fig. 3.4 Basic contention resolution scheme of IEEE 802.11

The DIFS[21] shown in the diagram is the minimum interval between two successive transmissions. The NAV[22] shown in the diagram represents explicit information on the length of the current transmission(s). In this scheme, prioritization of medium access in support of data flow Quality of Service is provided for by different minimum lengths of the back-off window.

The NAV improves the accuracy of the synchronization for channel access, but its usefulness is limited to homogeneous networks in which all stations conform to the same protocol standard. However, the IEEE 802.11 standard contains a second mechanism that operates on the basis of energy sensing only and, therefore, it can be used in general license exempt spectrum that is not application specific. Because energy detection is less reliable, it requires a higher threshold than (modulated) signal detection and therefore it is less efficient much like the energy sensing of EY-NPMA.

3.4.5 Setting the Channel Sensing Threshold

All methods of channel sharing that use distributed control but go beyond the "hope for the best" approach of the Aloha scheme, require some form of information distribution to the parties involved. This may involve implicit signaling (e.g. simple transmission) or explicit signaling as in the two wireless LAN protocols described above. In both protocol designs, it is assumed that all participants in the channel access are able to receive each other's signals correctly. To some extent, this is assured through the setting of a "channel busy threshold" above but near the minimum useful sensitivity of the receiver. Although this is beneficial for avoiding collisions with other transmitters, a very low threshold will reduce the opportunities of the device to transmit, even if it could cause no interference. Recall that interference depends on two things: the power output of the interferer and the SIR margin of the victim(s). Since the latter is not known, the setting of the channel busy threshold can only take into account the power output of the device's own transmitter. This is one of the factors limiting the efficiency of medium access schemes that depend on sensing.

[21] Distributed contention resolution Inter Frame Space.

[22] Network Allocation Vector.

3.4.6 Recovery from Channel Access Failure

All listen-before-talk protocols allow for a certain degree of channel access failure on an initial attempt, and they provide rules for recovery from such events. The IEEE 802.11 medium access protocol requires an exponential increase of the contention window with repeated failures (inferred from lack of an Ack reply). The EY-NPMA protocol does not specify this, but leaves it to the system's designer. This is a good choice since it avoids enforcing a solution that includes the potential of total starvation of some nodes. Under the rules of the IEEE 802.11 standard, certain conditions, e.g. interference or decoupling can cause transmission failures that, under the exponential back-off requirement, lead to starvation.

3.4.7 Impacts of Asymmetrical Coupling

In many cases, such as home networks, there are no hidden nodes and the protocol works as expected or it works at a level that the user does not notice any degradation.

However, as network speeds increase and wireless devices become more ubiquitous, this ideal state will not necessarily apply in general.

In a 2005 paper, Garetto[23] e.a. showed conclusively that there are cases in which sensing-based channel sharing fails. In each of these cases, it is possible that one member of a transmitter-receiver pair does not see both members of another pair. These configurations may occur when there is an obstacle or wall between some of the members of different pairs. Because of this asymmetry, the "uncoupled" stations may at any time start a transmission that can affect the exchanges of the other pair. The interaction between propagation asymmetries and medium access protocol is further explored in Chap. 7, Sect. 7.5, *Hidden Nodes and other Asymmetries*.

3.5 Summary of Sharing Dimensions

The preceding sections show that sharing mechanisms can be described as operating in one or more dimensions or parameter spaces – frequency, information, space, or time. These are all orthogonal and can be combined in various ways to suit a given application. The spatial parameter space has its three dimensions of its own – which further complicates the matter by adding another degree of complexity. Theoretically, there are many ways of combining the use of different dimensions; but in practice, certain combinations make more sense than others, in meeting the requirements of a given application or service. However, the analysis of the interactions in the unwanted domain always has to consider the full range of interaction possibilities.

---

[23] See Garetto et al. [52].

Sharing in frequency isolates the sharing participants in pairs or groups and, in case of the latter, a second sharing dimension is needed. Cellular CDMA systems are an example of a two-dimensional sharing system. Another example is sharing in frequency and time as prevalent in cellular TDMA system – regardless whether they use FDD or TDD.

Sharing in space is prevalent in fixed link systems that rely on highly directional antennas to achieve a high degree of spectrum sharing. Radar systems are another example of space sharing systems that can share a given frequency without much interference. Both rely heavily on very high gain antennas with narrow beams.

Sharing in time is prevalent in wideband systems like wireless LANs — its drawbacks of reduced efficiency and potential instability are more than offset by the ability to handle large changes in instantaneous throughput.

In this chapter, only some of the advantages and disadvantages of certain spectrum sharing methods were given. In the following chapter, specific sharing mechanisms and arrangements will be considered in more detail.

Chapter 4
Modes and Means of Spectrum Sharing

The preceding chapter mentioned spectrum sharing many times and in some cases, some parameters and procedures. In order to further delineate our subject – dynamic sharing of license exempt spectrum – various ways of spectrum sharing must be described.

4.1 Overview

As described in Chap. 1, spectrum sharing is defined as the condition in which – or the process by which – certain frequencies are used in a way that does not exclude others from using those same frequencies or range of frequencies. If the only relevant parameters are those that describe the static channel access criteria such as an RF power level, we speak of passive sharing; otherwise, we speak of active sharing.

Because of practical limitations such a non-perfect filters, radio systems have two domains of interaction with their environment: their operating frequency and the frequencies adjacent to the operating frequency. Where sender and intended receiver use the same frequency, we speak of on-channel interaction. However, all transmitters produce unwanted emissions of some kind, which include spurious[1] emissions, such as the harmonics of the transmitted frequency, as well as out-of-band emissions[2] that result from the modulation process. These can affect other than the intended receivers. The latter may have sensitivity profiles that extend well beyond the wanted frequency channel(s). This subject is described in more detail in Chap. 5.

[1]From ITU-R, Radio Regulations, Article 1, item 1.145: *spurious emission*: Emission on a frequency or frequencies which are outside the necessary bandwidth and the level of which may be reduced without affecting the corresponding transmission of information.

[2]From ITU-R, Radio Regulations, Article 1, item 1.144: *out-of-band emission:* Emission on a frequency or frequencies immediately outside the *necessary bandwidth* which results from the modulation process, but excluding *spurious emission.*

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_4, © Springer Science+Business Media B.V. 2011

Table 4.1 Interactions between transmitters and receivers

	Transmitter operating frequency	
Receiver operating frequency	f_0	$f_0+\Delta f$
f_0	On-channel interaction	On-channel interference
$f_0+\Delta f$	Off-channel interference	Off-channel interference

In case the interaction is unwanted we speak of on-channel interference; where either the transmitter or the receiver frequency overlap with the adjacent frequency of another receiver or transmitter, we speak of off-channel interference. See Table 4.1, which applies to FDD[3] as well as TDD[4] systems. The delta that defines the difference between on-channel and off-channel interference sensitivity is receiver specific.

Per the above definition, successful spectrum sharing means that on-channel and off-channel interference do not occur or are below certain agreed or prescribed levels. These levels are specific for the receiver, i.e. the victim system: at some level of interference, one type of receiver may be able to operate normally, whereas another may not be able to do so. Much depends on the receiver's required SNIR.

4.1.1 Vertical and Horizontal Spectrum Sharing

From a regulatory perspective, two forms of spectrum sharing can be distinguished: horizontal sharing and vertical sharing.

Vertical refers to services or users having a different regulatory status. This is the case when a service has to share spectrum with a service having a primary allocation. Two examples are:

(a) the Amateur Service, which has typically secondary status in spectrum allocated on a primary basis to, for example, the Radiolocation Service, and
(b) the Mobile Service in the 5 GHz band, which has co-primary allocation with the Radiolocation Service, but has to operate on a non-interference/non-protected basis.

Horizontal refers to "having the same regulatory status." This applies between devices in a license exempt band, but also between devices in other bands such as the Private Mobile Radio band: the regulator sets RF power limits and frequency limits and, in the case of PMR, certain protocol criteria and that is all. The practice of sharing is left to the equipment designers and, notably in the case of PMR, to the users.

Another interesting case of horizontal sharing is sharing between VSATs downlink operations and the earthbound WiMAX systems in the 3.4–3.6 GHz band. Historically, this band is an "extension" band which is to be used if and when the

[3]Frequency Division Duplex – stations can simultaneously send and receive – on different frequencies.

[4]Time Division Duplex – station use the same frequency alternately for sending and receiving.

VSAT community would need more than its initial allocation in the 3.6–4.2 GHz VSAT downlink band.[5] The WiMAX community has pointed out that this extension band is underused in many places (thanks in no small part to the vast capacity of the transatlantic fibre optic cables). Provided WiMAX systems take certain pre-cautions, sharing this band is perfectly possible. Many regulators have recognized this and have offered licenses to prospective WiMAX operators. The technical details of this interesting case are outside the scope of this book.

4.1.2 Managed Sharing

From an operational perspective, spectrum sharing occurs in two main forms: managed sharing and unmanaged sharing. In managed spectrum sharing, the parties involved have the same regulatory status – both are primary. Their operating frequencies and transmitter locations are carefully planned so as to avoid significant interference. Managed Sharing involves at least two licensees, sometimes more. The prime example of managed sharing are the cellular mobile phone systems: they use CDMA,[6] TDMA[7] or W-CDMA[8] medium access and have in common that they operate in FDD mode; uplink and downlink frequencies are separated such that stations may transmit and receive at the same time. The frequency separation required between uplink and downlink is appreciable – as much as 7–10% of the operating frequency. In such systems, interference occurs only between signals from adjacent cells and therefore the combination of a few frequency channels and careful cell separation can easily assure that the system's SNIR is not affected by spectrum sharing. Managed sharing is equally important for Fixed Link systems – like cellular systems, they operate near the limits of technical feasibility in order to achieve their high levels of performance. Other examples of managed sharing are the assignment of frequencies to radar systems such that unnecessary interference is avoided. However, the scope and accuracy of such arrangements are far less than what is needed for cellular mobile systems.

Unmanaged sharing is the rule rather than the exception in the license exempt domain: it may involve a vertical regulatory relationship between the sharing parties – as in the case of radars and wireless LANs in the 5 GHz band, or it may involve a horizontal regulatory relationship – as in the case of users of the ISM bands. However, there are many frequency bands which are allocated to applications ("services" in the ITU-R parlance), with sharing requirements left to the users of the so allocated spectrum. In the following section, we are concerned with unmanaged sharing.

[5]The uplink band is at 5.8–6.4 GHz.

[6]Code Division Multiple Access, see Chap. 3, Sect. 3.2.

[7]Time Division Multiple Access – it is variation of TDD mode.

[8]Wide band CDMA – the hybrid access mode of the 3G cellular systems.

4.1.3 Modes of Sharing

Spectrum sharing may done passively – in this case, users are oblivious to the existence of others – or actively – in which they have to be aware of and interact with other users.

4.1.3.1 Passive Sharing

Passive sharing refers to the absence of real-time measures to facilitate spectrum sharing. This means that on-channel interaction is well below the receiver's required SNIR and the adjacent channel interference levels are below some acceptable level. An obvious case of passive sharing occurs when two communities are separated physically, e.g. by a hill or mountain range that prevents transmissions in one community interfering with the receivers of the other community. Here, passive sharing is based on spatial separation. A similar situation may exist in the time dimension: a set of frequencies may be used in daytime by private communications networks; at night, radio amateurs may be using the same frequencies without much chance of interfering with a nearby private network.

A less obvious case is the use of signal coding to achieve a level of robustness against on channel interaction and interference such that the presence of interfering signal below a certain level does not affect the performance of the receiver. The best known example is the CDMA cellular system.[9] A more generic case is the use of certain robust protocols by multiple users of a frequency range who may be in communications or interference range with all or with some members of the user community. An example of such a protocol is the Aloha protocol,[10] which boils down to: "transmit at any time, repeat if no acknowledgement." There is no means of coordinating transmissions and the resulting loss of capacity is a function of traffic volume. The on-channel interaction in this case is present, but it is far from constant: at low traffic volumes Aloha (Pure or Slotted) can be a perfectly adequate means of sharing spectrum – even though its mode of sharing is passive. This subject is further treated in more detail in Chap. 6, Sect. 6.1.

Another case of passive sharing is that of Wi-Fi and Bluetooth devices in the same notebook computer: the Wi-Fi device provides the link to the access point, the Bluetooth device may provide a link to a smartphone or headset. Because the standards governing the implementation of these technologies do not take into account each other – or any other standard – the two devices have no way of adjusting to each other's activity and the result is an predictable interference. Unless continuous separation in frequency[11] is possible, interference can be avoided only by making

[9] Known as IS-95 and IS-2000 in the US.

[10] See Abramson [8].

[11] Bluetooth can adjusts its frequencies of operation such that it avoids overlap with the frequencies used by a wireless LAN device.

the operation of one device contingent upon the other being inactive or disabled – in other words, active sharing at the application level would be required.

4.1.3.2 Active Sharing

This term refers to real-time measures taken to avoid interference. The motivation for such measures can be regulatory – i.e. the measures are required – but also self-interest: avoiding interference typically means not receiving interference either; and therefore, performance is optimized. In the terms of Table 4.1, Active sharing strives to avoid on-channel interaction dynamically. The avoidance can be implemented in frequency, time or coding space and can take various forms, depending on the type of system and its transmission protocol. The procedures by which users avoid causing interference is sometimes referred to as a “spectrum etiquette.” See below in Sect. 4.3.

In general, active spectrum sharing requires that a transmitter determines if its intended channel of operation is in use by other systems. In a listen-before-talk (LBT) etiquette, the transmitter obtains the required information autonomously by listening on the intended channel. Alternatively, active sharing can be dependent, i.e. based on obtaining permission from another system. The permission-based approach can be implemented using another radio-based system or a separate control channel; it may be based on a spectrum usage data base.

In the autonomous case, the transmitter listens for signals from other systems whose transmissions it seeks to avoid. In case of avoiding interference in time, the transmitter only determines the RF energy level in its channel; in the case of code avoidance, it has to listen for transmissions with its own intended code pattern. In practice, there is not a clear boundary between these two modes of signal detection: simple energy detection is less precise and less selective, but true signal detection takes time. The lack of precision may cause unnecessary deferral, whereas higher sensitivity may cause deferral to sources that are not relevant. Therefore, some systems combine the two methods. A case in point is the IEEE 802.11-2007 standard, which has both a “carrier detect” function and an optional energy detect function that serve its channel access determination. The former listens for the training sequence of transmission frames and is 15 dB more sensitive than the latter. Therefore, systems based on this standard will defer to their own kind more readily than to other systems which do not use its particular training sequence.

Deferral can be done on timescales that range from days (or more) to microseconds – the technical necessity depends on the system’s application and protocol, but regulatory requirements may override this. There is difference between (real-time) applications requiring instant channel access and non-real-time applications in which channel access may be deferred for a limited time. An example of the former is a remote control system for a robot device, an example of the latter is a data link for file-based data transfer. A channel selection mechanism may be needed for the remote control systems, whereas the file transfer allows deferral to other users on a per transmission basis.

Regulatory requirements may enforce a channel selection mechanism regardless of the nature of the transmission protocol of the interferer. A case in point is the DFS requirement[12] for devices that operate in spectrum shared with radar systems in the 5 GHz range. This subject is treated in more detail in Chap. 9, Sect. 9.1.

4.2 Techniques for Spectrum Sharing

4.2.1 Spread Spectrum Techniques

Spread spectrum techniques use the frequency dimension to share with other systems – by using a wide range of frequencies, the interference potential is reduced. Spread spectrum techniques come in two basic flavors:

(a) Direct Sequence Spread Spectrum (DSSS) – which spreads in the frequency domain only and which is comparable to CDMA,[13] and
(b) Frequency Hopping Spread Spectrum (FHSS) – which spreads the use of a given frequency in time.

In DSSS, the carrier is spread over all or over a large part of the operating bandwidth. This is achieved by modulating the carrier by bit sequences that represent a given symbol. A symbol may correspond to one or more bits, e.g. a 1 or a 110. The redundancy bit sequences are known as chip sequences, which are typically pseudo random numbers (PNs). Transmitter and receiver have to use the same PNs. The ratio between the data symbols and the chips is known as the chip rate which equals the processing gain: a chip rate of 100 gives 20 dB processing gain. This gain allows the receiver to recover signals buried in channel noise or interference: e.g. for a symbol rate of 1 mega symbol per second and one data bit per symbol, a few dB of SNR are needed. However, spreading the signal to 22 MHz allows the receiver to recover it from 10 dB or more below the receiver's noise floor. This small processing gain was made use of in initial the IEEE 802.11 standard for wireless LANs, but only as an interference mitigation mechanism – it is not sufficient to truly share spectrum. The primary mechanism for spectrum sharing in that standard is collision avoidance, which operates in the time domain and requires sensing.

In FHSS, the carrier is switched at a certain interval to some frequency slot of the operating frequency range. The order of these frequency slots is typically a pseudo random sequence. The number of frequencies needed is function of the number of users sharing the frequency band and the degree of separation required: unless hop sequences are synchronized and parallel, the rate at which overlap between two sequences occurs depends on the number of hop sequences available. If each hop sequence is n hops and there are m users, the probability of a given frequency slot

[12] "Dynamic Frequency Selection", see also Chap. 9.

[13] See Chap. 3, Sect. 3.2.

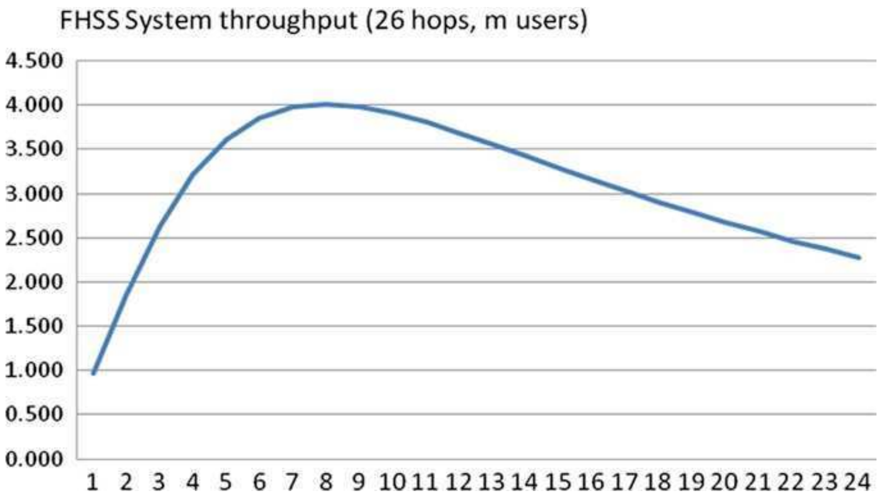

Fig. 4.1 IEEE 802.11 FHSS network throughput in Mb/s vs. the number of active users

being free for a given user is (1 – 1/n)m. This applies for common slot lengths and full synchronization between the hoppers. The aggregate throughput on that set of hops is roughly as given by the curve in Fig. 4.1, which is generated by the formula:

$$R(m) = \frac{\mathrm{m} \times \mathrm{PER}}{1 + \mathrm{rm}} \tag{4.1}$$

in which: m is the number of users, PER = (1–1/n)m in which is n the number of hops and rm is the retransmission rate for a 10% or lower frame failure rate.

4.2.2 *Directional Antennas*

Directional antennas are typically used for extending the range of a radio systems; the potential benefits for spectrum sharing are frequently ignored. A good example is the use of directional antennas for point-to-point links operating in the 2.4 GHz band and using Wi-Fi transmitters and receivers: regulations that limit radiated power prevent users from obtaining maximum benefit from their equipment.[14]

As shown in Chap. 5, Sect. 5.1.3, directional antennas facilitate spectrum sharing by reducing the “interference footprint” of the transmitter. The improvement depends on the antenna gain, as well as on the environment; it becomes noticeable at antenna gains of e.g. 6 dB and more for heavily cluttered environments and 20 dB

[14]Whether these restrictions invite illegal practices is another matter which falls outside the scope of this book.

and higher for open environments. Clearly, the higher gain antennas offer better spectrum sharing possibilities.

The secondary effect of higher gain antennas is that the level of unwanted emissions is also increased over the footprint of the antenna. However, the antenna itself has certain frequency selective properties that may reduce this increase. Regardless of antenna selectivity, the area reduction factor of the antenna applies. For a given power output, the probability of interference, also in the out-of-band domain, is less for directional antennas than it is for omni-directional antennas.

A general drawback of the use of fixed directional antennas is that their use is not flexible: notably high gain antennas are difficult to point and align so as to achieve optimum results. Steerable directional antennas avoid some of this inflexibility, but at the cost of increased complexity (physical/electrical steering) or power consumption (electronic steering). As technology progresses, the difference between fixed and steerable antennas may become less pronounced. To some extent, this is already happening: practical electronically steerable 60 GHz antennas have been developed and are finding acceptance in the market.[15,16]

4.2.3 Medium Sensing

Sensing the presence of other spectrum users provides a basis for a variety of sharing approaches that use either the frequency domain or the time domain or a combination of these. Medium sensing clearly belongs to the scope of active spectrum sharing, and it can be applied to sharing involving similar systems as well as sharing involving dissimilar systems. In the case of the former, the sensing mechanism can be made selective as well as sensitive: recognizing signals with certain properties is easier than sensing energy.

Sensing has drawbacks, notably the general uncertainty associated with using distance as the sharing dimension. As explained in Sect. 3.3.1, this applies to all cases and modes of spectrum sensing, notably in open propagation environments. Therefore, sensing functions best in environments with a high pathloss exponent but with limited shadowing – typically found in rooms and other enclosed spaces. The success of the base-line Wi-Fi technology based on the IEEE 802.11 standard, as well as its weak points, demonstrate this.

Another drawback that plays a role in sharing between dissimilar systems is bandwidth asymmetry: see the interference behavior of narrowband and wideband systems as explained in Sect. 4.1. Table 4.2 gives an example case: a 1 MHz wide narrowband system and a 20 MHz wide wideband system with a similar spectral efficiency of .5 bit/Hz.

The interference power of the narrowband signal at the distance at which it is able to detect the wideband system is 2.2 dB above the minimum discernable signal level of the wideband system. This causes range loss in the wideband system.

[15] See e.g. Tanaka and Ohira [119].

[16] See for a practical example Gilbert [57].

Table 4.2 Impact of asymmetrical bandwidth on interference impact

	Narrowband system	Wideband system
Frequency of operation (GHz)	2.4	2.4
Bandwidth (MHz)	1	20
Power limit (dBm)	20	20
Data rate (kb/s)	500	10,000
Required BER	0.00001	0.00001
SIR at data & BER (dB)	8	12
Receiver implementation margin (dB)	6	6
MUS	−100	−83
Detection Threshold (=MUS + SIR + 6 dB)	−102	−89
Fading & Shadowing margin (dB)	10	10
Required Rx signal (dB)	−90	−73
Pathloss exponent	3.3	3.3
1 m pathloss (dB)	40.1	40.1
Operating link budget	120.0	103.0
System operating range	264.1	80.6
Interference detection link budget (dB)	112.8	109.0
Detection distance for the other system (m)	159.7	122.5
Interference power at detection distance, seen by the other system (dB)	−92.8	−102.0
Interference power relative to noise floor + implementation margin of the other system (dB)	2.2	−10.0

On the other hand, the wideband signal at the range at which the wideband system detects the narrowband signal is 10 dB below the latter's minimum discernable signal level. This difference of 12.2 dB between the interference impact of these two systems implies that, in a mixed environment with a propagation exponent of 3.3, the effective coverage area of the narrowband system is (212.2/10)2 = 5.27 times[17] larger than the effective coverage area of the wideband system. Theoretically, this increases to a factor 16 in free space conditions.[18]

4.2.4 Database Look-up

Database look-up as a means to share spectrum, which is partially used geographically and frequency wise, has been proposed as the best mechanism for sharing between wireless microphones and license exempt devices in the TV White Space bands and it has been accepted by the FCC.[19] The UK's Ofcom is expected to follow a similar route to facilitate sharing of underused spectrum in the TV bands.

[17] A pathloss exponent of 3.3 is equivalent to 10 dB pathloss per octave ($10\log 2^{3.3}$ ~ 10).

[18] A pathloss exponent of 2 is equivalent to 6 dB pathloss per octave ($10\log 2^{2}$ ~ 6).

[19] These are frequencies between 40 and 600 MHz that become available as broadcasters transition from analog to digital TV. See FCC 2008, Second Report and Order 08–260, ET docket No. 04-186.

Data base look-up and the attendant registration of spectrum users have been used for many years, also in cross-border situations. An example of the former in the US is the registration of 70/90 GHz point-to-point links and the registration data base of Ireland for 5.8 GHz systems; the Dutch/German database for fixed links is an example of the latter type of use. These examples all involve static, non-real time operation: new sites can be checked out before installation and, in case of interference claims, the data base provides a reference to the owners of the systems concerned.

When a data base is used to support spectrum sharing in real-time, as in the case of TVWS spectrum, the database entries cover stations and other protected systems at the edges of the protection area. For TV transmitters, the edge is determined by a regulatory protection contour. In theory, the database could record locations that are known to be safe for use by TVWS devices, in the sense that interference into TV receivers has not been reported from these locations. Assuming TVWS users report their use of locations and frequencies, the data base would eventually map the separation distances needed between TVWS devices and DVT receivers to avoid interference. The security aspects and the legal implications of database driven spectrum sharing need not concern us here. However, there are significant technical issues to consider. Although data base approach may appear to avoid the downsides and pitfalls of spectrum sensing, it does not. Protection contours are always imprecise and require “padding” to assure a modicum of protection of the incumbents. Filling the data base with “reality data” is at best a trial and error operation that depends on propagation conditions, including the weather and, rather obviously, the presence of active TV users at the time a TVWS device is installed or operated. Conditions will change with time and so will the effective interference protection boundary.

4.3 Spectrum Sharing Techniques for Dissimilar Systems

The term “dissimilar” is used here to denote significant variation in essential properties.[20] In other words, the systems concerned are to be treated as the proverbial apples and oranges when it comes to spectrum usage properties. The dissimilar systems involved in a sharing regime may have the same or a different regulatory status – in the first case, one speaks of horizontal sharing; in the second case, one speaks of vertical sharing.

4.3.1 Overview

Spectrum sharing between dissimilar systems is the rule in license exempt spectrum: any device, type modulation, or access protocol is possible within the bounds

[20]Examples of essential properties in this context are modulation scheme and medium access protocol.

set by the regulator. The 2.4 GHz ISM band is the best known example of spectrum used by a wide variety of systems, ranging from hearing aids to outdoor point-to-point links. With time, a similar situation may develop in the 5 GHz band. RF power outputs range from a few milliwatts (e.g. Bluetooth) to a full Watt (e.g. a 2.4 GHz point to point link).

Spectrum sharing rules for dissimilar systems are frequently referred to as a spectrum etiquette. That term is appropriate: just like an etiquette among people serves to canalize interactions in generally preferred patterns, a spectrum etiquette is a set of simple rules that aim to avoid – or at least reduce the probability of – spectrum users colliding. And, just like human etiquettes, a spectrum etiquette avoids the use of explicit information exchange. Thus, a spectrum etiquette can be used among systems with different characteristics. It typically requires device to listen for the activity of other devices and to decide on spectrum access based on the results. Therefore, spectrum etiquettes are examples of active spectrum sharing mechanisms. Because such etiquettes generally assume some homogeneity of spectrum utilization parameters, they belong to the category of horizontal spectrum sharing.

Spectrum sharing between dissimilar systems may occur in any piece of spectrum, regardless its allocation status. A prominent example is the sharing arrangement in the 5 GHz band between the radars operating as a Radiolocation Service – which has primary status – and wireless LANs operating as a Mobile Service. The latter are required not to cause interference into the radar systems and, therefore, they have to detect their presence and avoid using their frequencies.

Ultra-Wide-Band systems are another example of a case of sharing between dissimilar systems: its transmissions span a vast range of frequencies that overlap with a number of bands allocated on a primary basis to different services. For each, different sharing criteria apply; this is reflected by the regulatory requirements imposed on Ultra-Wide-Band transmitters.

Historically, little has been achieved in spectrum sharing between dissimilar systems or technologies: each technology has its own preferred modulation schemes and channel access techniques. Notwithstanding the good intentions of a forum like the IEEE 802.19[21] Committee, no progress has been made in this area. Nonetheless, this form of spectrum sharing is the most challenging and potentially the most rewarding; and therefore, this book addresses that subject in some detail, both from a theoretical point of view and from a practical point of view.

4.3.2 Issues with Sharing of License Exempt Spectrum

The issues that dominate this form of spectrum sharing are many – if only because the absence of design constraints allows any form of spectrum usage that fits the applicable power limits. Therefore, designers are free to use any or all of the spectrum

[21] See Berlemann and Mangold [24], p. 45.

sharing dimensions given in Chap. 3: frequency, information, space, and time. This degree of freedom is a two-edged sword: it allows designers to follow their whims, but those whims can prove very detrimental to others.[22]

Each of the sharing dimensions has its own properties with regard to the costs and benefits of spectrum sharing.

4.3.2.1 Spectrum Sharing Using the Information Dimension

The simplest of the four dimensions – the information dimension – is also simple in its application and its consequences: the more redundancy is added to the signal, the lower its power density and the lower its interference potential. Conversely, that same redundancy reduces the capacity as well as the interference sensitivity of the system and it increases its robustness correspondingly. See Sect. 3.2. Therefore, using this dimension to share spectrum with others has the potential to share "blindly," i.e. passively. Given adequate redundancy, the need for awareness of others is avoided because, for the applicable RF power limits, the power density drops and that reduces the interference impact; thus, the use of other dimensions to improve sharing behavior and/or impact may not be necessary. An example is spread spectrum system that uses a large spreading ratio, e.g. 40. For the same power output, its power density is 16 dB less than a non-spread system of similar throughput and its interference range is significantly reduced. On the other hand, those 16 dBs make the system very robust against interference.

4.3.2.2 Spectrum Sharing Using the Frequency Dimension

Next in complexity is the frequency dimension. In the license exempt world, there are no channels or other common factors to leverage and, therefore, some form of active sharing is needed: spectrum usage has to be preceded by a check on the status of the frequency to be used: either by sensing, by listening to a control channel[23] or by a data base look-up. The use of a common control channel implies some form of shared properties or similar capabilities and which are typically absent in license exempt systems. Similarly, the data base method assumes a degree of common capabilities absent in systems deployed in license exempt spectrum. That leaves sensing and its attendant issues. As explained in Sect. 4.2.3, sensing other spectrum users has inherent limitations in terms of accuracy.

---

[22] This is the main reason why spectrum regulators typically limit the output power in frequency bands that are license exempt. Since propagation affects signal strength and therefore interference power exponentially, limiting RF power will in many cases limit the number of potential victims.

[23] Common Control Channels have been proposed many times, but few implementations exist – This form of sharing may well be attractive only in case all deployed devices are under the same owner or authority.

The implications for a designer are that a choice must be made between the protection of others and the capacity requirements of the own system. Unless constrained by regulatory requirements, that choice will favor the own system at the expense of other users of that spectrum. Therefore, regulatory constraints are required to avoid that some technologies dominate the use of license exempt spectrum.

4.3.2.3 License Exempt Spectrum Sharing Using the Space Dimension

As noted in Sect. 3.3.2 Directional Radiators, sharing in the space dimension using directional, high gain antennas is very efficient: in theory, many users can share the same physical space because the probability of interference is roughly proportional to the square of the antenna angle. The probability of interfering with randomly located non-directional systems is a fraction of that probability for non-directional transmitters. See also Sect. 4.3.4 on regulatory considerations. Deployment of directional antennas has to contend with the inherent inflexibility of maintaining pointing accuracy, which gets more difficult with increasing antenna directivity. Alternatively, some form of "target tracking" may be applied that allows transmitter and/or receiver mobility without losing the advantage of high spectral efficiency and low interference probability. The emergence of low-cost steerable antennas – mostly at very high frequencies – has eased this situation considerably. A good example is Sibeam's beamforming WirelessHD system.[24]

At lower frequencies, beamforming is being deployed to realize some of the advantages of spatial spectrum re-use. A case in point is MIMO/OFDM PHY layer of WiMAX and 4G and LTE systems. The specifications provide for beam steering to improve cell range and capacity. The same technology could be applied in case of license exempt systems.

4.3.2.4 License Exempt Spectrum Sharing Using the Time Dimension

Wireless systems using CSMA/CA are widely deployed and very successful even though the CSMA/CA protocol requires full "visibility" of all participating nodes. Hidden nodes can lead to starvation issues. In many cases, such as home networks, there are no hidden nodes and the protocol works as expected or it works at a level that the user does not notice any degradation.[25] However, as network speeds increase and wireless devices become more ubiquitous, this ideal state will not necessarily

[24] See for more information http://www.sibeam.com/

[25] In home networks that depend on an ADSL or cable gateway, the limiting factor is typically the gateway, not the wireless LAN's channel utilization.

apply in general. In a 2005 paper,[26] Ed Knightly e.a. showed conclusively that there are cases in which sensing-based channel sharing fails. This issue is discussed in more detail in Chap. 7, Sects. 7.4 and 7.5.

4.3.3 Cognitive Radio Techniques

Since the introduction of this concept by Mitola,[27] much work in this area has been done by various research organizations such as the DARPA XG project[28] and other fora such as the European Telecommunications Standards Institute (ETSI). In the Wireless World Research Forum of 27 October 2003, Preston Marshall, Program Manager of DARPA XG Program, said, "The Primary Product of the XG Program is Not a New Radio, but a Set of Advanced Technologies for Dynamic Spectrum Access." The assumption behind this idea is that by combining a variety of technologies, radios can make autonomous decisions about the spectrum to use – in a given location and at a given time – that best matches the user's communications requirements. The tools and technologies under consideration in the FCC's Report & Order of 2005 on Cognitive Radio include[29]:

1. Frequency Agility – the ability of a radio to change its operating frequency to optimize use under certain conditions.
2. Dynamic Frequency Selection (DFS) – the ability to sense signals from other nearby transmitters in an effort to choose an optimum operating environment.
3. Adaptive Modulation – the ability to modify transmission characteristics and waveforms to exploit opportunities to use spectrum.
4. Transmit Power Control (TPC) – to permit transmission at full power limits when necessary, but constrain the transmitter power to a lower level to allow greater sharing of spectrum when higher power operation is not necessary.
5. Location Awareness – the ability for a device to determine its location and the location of other transmitters, and first determine whether it is permissible to transmit at all, then to select the appropriate operating parameters such as the power and frequency allowed at its location.
6. Negotiated Use – a cognitive radio could incorporate a mechanism that would enable sharing of spectrum under the terms of a prearranged agreement between a licensee and a third party. Cognitive radios may eventually enable parties to

---

[26] See Garetto et al. [52], August 28–September 2, 2005, Cologne, Germany.

[27] The idea of cognitive radio was first presented officially by Joseph Mitola III in a seminar at KTH, The Royal Institute of Technology, in 1998, published later in an article by Mitola and Gerald Q. Maguire, Jr in 1999. See also Mitola's thesis: *Cognitive Radio: Model based competence for software radios* (http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-970).

[28] See e.g. Murthy [98].

[29] See FCC 05-57 – Facilitating Opportunities for Flexible, Efficient, and Reliable Spectrum Use Employing Cognitive Radio Technologies, page 5.

negotiate for spectrum use on an ad-hoc or real-time basis, without the need for prior agreements between all parties.

The military background to the above is clearly visible: resources and complexity considerations are not the primary concern. One might add that practical considerations also got little attention – as did the physical constraints that apply to the accuracy of spectrum decisions reached by cognitive radios. In ETSI Technical Report TR102 799, a detailed analysis is given of all the factors that play a role in dynamic spectrum access – seen from the perspective of a digital wireless microphone system that uses cognitive techniques to find a suitable operating frequency in the VHF/UHF bands. It concludes that there are many challenges in realizing such a system. Notably, sensing is – correctly – stated to be inadequate as basis for finding a free channel or for populating a spectrum utilization data base, the reasons including hidden nodes and propagation conditions that even locally vary widely. This assessment indicates that the concept of Cognitive Radio is not a panacea, but, if applied with care and understanding, an approach that will provide excellent results.

The above should not cause Cognitive Radio to be written off as yet another hype. The general concept is useful and its application does in some cases yield excellent results, notably in the case of the sharing of the 5 GHz band between wireless LAN and radar systems. More on this subject is given in Chap. 9, Sect. 9.1.

4.3.4 Regulatory Considerations: RF Power Limits

The main reason why spectrum regulators limit the RF output power in frequency bands that are license exempt is to reduce their interference potential. Propagation affects signal strength (and therefore interference power) exponentially, and limiting RF power will in many cases limit the number of potential victims without reducing the operating margins of license exempt devices to unacceptable levels. There are three ways to specify RF power limits: radiated power or EIRP,[30] RF output power and RF power density.

4.3.4.1 EIRP Power Limit

EIRP is an abstraction that gives the total radiated power as measured over a sphere surrounding a hypothetical perfect radiator. In practice, it is measured on the real device over a number of angles; the highest value must be below the given limit. EIRP power limits are favored by some regulatory authorities: it is simple to spec-

[30] Equivalent Isotropically Radiated Power – this measure gives the total power radiated equally in all directions by an ideal dipole antenna.

ify, it assures that the interference seen by other spectrum users is never more than the specified limit and it leaves designers all the freedom they may want.

Although this is good at first sight, the implications – as shown in Chap. 5, Sect. 5.2.3 – are negative from the viewpoint of spectral efficiency: with directional antennas, more systems can share the same geographical and frequency space than with non-directional antennas. Therefore, EIRP limits tend to leave spectrum unused. In some cases, the regulator has recognized this. For example, the FCC's Part 15 rules which specify an RF power limit Ptx and allow for antenna gain to be taken into account: the EIRP is limited to PEIRP = Ptx + 6 dB - (G-6)/3 – in which G is the highest antenna gain. This rule is still conservative, but it does recognize the benefit of directional antennas.

4.3.4.2 Total RF Power Limit

A total RF power output limit simply specifies the maximum RF power that the transmitter may produce while disregarding the effects of bandwidth and antenna directivity. Specification of an RF power limit has the advantage that it is neutral with regard to other spectrum users: wider bandwidth means a reduced power density and, therefore, reduced spatial interference potential.

Further, an RF power output limit is ideal from a system's design point of view as it allows the designer to optimize the combination of RF power output and antenna gain patterns to match a given operational profile – which may call for long range or a large coverage area or high data rate – or some mixture of thereof.

A possible downside is the increased complexity of the rule making: the antenna gain must equate to azimuthal directivity – at least above 6 dB. This is shown in Chap. 5, Sect. 5.2.3.

4.3.4.3 RF Power Density Limit

The above type of limits do not consider the bandwidth of the emitted signal. For a given RF power output level, the interference potential of a signal increases as its bandwidth decreases –and vice versa. A well-known example is Ultra-Wide-Band technology in which the RF power is spread over such large bandwidths that the interference power seen by "normal" receivers approaches the background noise level. However, the reverse is also true and therefore, in order to protect other spectrum users from very high power density emissions from other users, a RF power density limit is useful.

4.3.4.4 Technology Neutral Power Limit

These considerations lead towards the conclusion that regulatory limits for dissimilar systems that are technology neutral are best given in terms of transmitter power

spectral density and regardless of antenna gain. However, without a limit on the total amount of power allowed, the only de facto limit on transmitter power output would be cost and/or energy consumption. This would remove any incentive to design efficient radio systems.

Therefore, a regulatory limit that applies dissimilar systems operating on the basis of unmanaged sharing is best given as the tuple [P],[PSD].

4.4 Spectrum Sharing Between Similar Systems

Spectrum sharing between systems that have similar spectrum usage properties is technically a less demanding subject than spectrum sharing between dissimilar systems. Notably active sharing is more easily analyzed, defined and executed if systems are alike in their essential properties. Similar RF power levels facilitate setting interference margins, similar modulation schemes facilitate detection of presence or activity and similar protocols facilitate defining rules that allow such systems to detect each other and to avoid collisions in the use of spectrum.

The most successful and most used forms of active sharing between similar systems are Bluetooth and Wi-Fi, even though the two are very different from each other. Both take note of the presence of other members of their own family and adjust their spectrum access accordingly.[31] Both are discussed in more detail in Chap. 6.

4.5 Summary

The preceding sections discuss various spectrum sharing arrangements, which are summarized in Table 4.3.

Note: vertical sharing between similar systems is possible in theory, but not applicable in practice. There is little interest in systems that are subject to non-interference requirements from other than their own kind, but operated by another player.

In general, passive sharing requires either very careful (and therefore costly) planning and installation – as in the case of sharing between primary – licensed – users. The alternative is to rely on wide margins to assure non-interference. This is notably true in the case of vertical sharing where it leads to inefficiencies in the allocation and the acquisition process or it wastes spectrum capacity. Since neither is very attractive, active sharing will be preferable if only because it has the potential to realize the maximum capacity in a given sharing arrangement. In addition, active sharing technology holds the promise of being flexible and adaptable to changes in conditions that a planned deployment never could adapt to.

---

[31] Some Bluetooth designs implement “adaptive hopping which avoids interference with – and from – wireless LAN devices and other spectrum users.

Table 4.3 Unmanaged spectrum sharing: arrangements and methods

	Similar systems		Dissimilar systems	
	Active	Passive	Active	Passive
Vertical sharing	Not applicable	Not applicable	Preferable – case specific rules e.g. DFS	Possible – e.g. channel presets, data base look-up
Horizontal sharing	Preferable – medium access protocol	Possible – e.g. channel presets, data base look-up	Preferable – spectrum sharing etiquette	Possible – e.g. channel presets, data base look-up

The promise of flexibility is carried to its extreme in the concept of Cognitive Radio. In a way, that concept is purported to do away with the implications of Table 4.1, because Cognitive Radio is supposed to be *the solution* to all situations and conditions of spectrum sharing. Throughout the analysis of spectrum sharing in the following chapters, that claim will be revisited.

Chapter 5
The Physics of Spectrum Sharing

5.1 Introduction

Shared spectrum use has demonstrated its potential with the technologies that today populate the ISM[1] bands at 900 MHz, 2.4 GHz and 5.8 GHz.[2] Based on these successes, making more spectrum available for shared use is assumed to provide relief from the purported spectrum scarcity that has many pundits up in arms against spectrum managers and regulatory rulemaking.

Spectrum designated for shared use, including the above ISM bands, is subject to certain rules. Such rules may limit users of a given frequency band to only one type of system or, as in the case of license exempt bands like the 2.4 GHz band, they impose little more than RF power restrictions. Notably in the ISM bands, communications devices may be exposed to interference from their own kind as well as from other interfering sources, such as microwave ovens, cordless phones, and radar systems operating in the same spectrum. The regulatory power output restrictions limit the potential damage caused by communications devices to each other and other spectrum users.

In shared spectrum, interference is the major factor affecting radio system performance in terms of throughput and quality of service. The most severe interference is typically co-channel interference caused by other wireless devices in the vicinity. In addition, there may be adjacent and alternate adjacent channel interference due to partially overlapping channel use, interference from non-compliant devices and from mis-configured or maliciously operated devices. In order to understand how interfering signals influence system performance, one has to look at the whole chain of elements involved in the transfer of information over a radio channel: the transmitter, the transmitting antenna, the radio channel itself, the receiving antenna and the receiver – as shown in Fig. 5.1. The transmitter includes a digital

[1] Industrial, Scientific and Medical applications.

[2] Formally, the 5.2 and 5.4 GHz bands are not license exempt in the same way as the ISM bands. These bands have been allocated by the ITU-R to the Mobile Service. But on a no-protection, no-interference basis.

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_5,

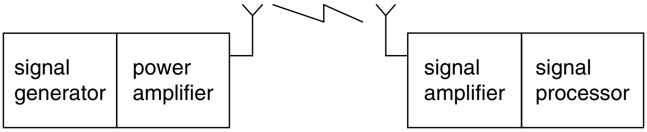

Fig. 5.1 A conventional (SISO) radio system

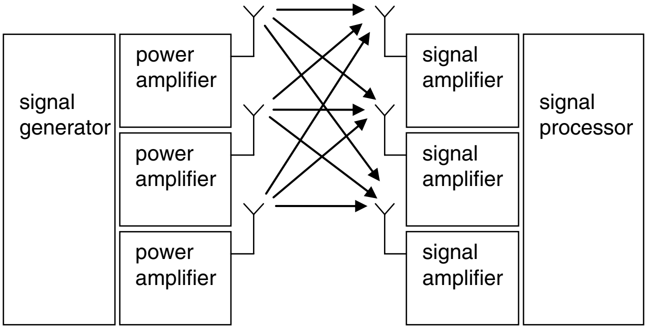

Fig. 5.2 A simplified MIMO system block diagram

signal processor (DSP) converting a (analogue or digital) source signal into a radio signal and an amplifier to boost the power of the radio signal, which is fed into the transmission antenna. The radio channel, the particular transmission path between sender and receiver, changes the signal in a number of ways, which include not only attenuation but also frequency selective fading, as well as adding background RF noise and interference. The receiver's antenna picks up the signal from the transmitter, possibly mixed with other signals and passes this to the receiver. The receiver performs signal processing to decode the received signal and extract the wanted information. Together with the propagation conditions, inherent noise and interference affect the operation of all those components.

A more complicated picture results from the use of channel models with multiple ports at each end. In practice, these ports correspond to antennas: input ports model transmit antennas, output ports model receive antennas. The "Multi I/O" variations cover cases such as the single input/single output case (SISO), the single input, multiple output case (SIMO), which includes the traditional receiver diversity, multiple input/single output case (MISO), which includes the traditional transmitter diversity scheme and, finally, the multiple input/multiple output case (MIMO), which includes diversity options at both ends of the channel. See Fig. 5.2. More detail is provided in Sect. 5.5.2 MIMO below.

Regardless of the type of channel choice, each set of channel input and output port has a maximum capacity that is given by the Shannon-Hartley theorem, which states that bandwidth and channel capacity are closely linked:

$$C - B \times log_2(1 + SNR) \tag{5.1}$$

Without a very large Signal to Noise Ratio (SNR), high capacity requires a large bandwidth and vice versa. This theorem applies regardless of the conditions: receiver that has to extract information from an interference rich channel can only succeed to the extent that the wanted signal exceeds the interference.[3] If the interference is noise-like, it can be used in Eq. 5.1 to determine the useful bandwidth.

The following sections address the impact of interference effects on signal transmission, signal propagation, and signal receiving respectively. The purpose is not to repeat textbook knowledge of basic radio technology, but to point to factors that affect spectrum sharing – positively or negatively.

5.2 Signal Transmission

Signal transmission consists of three processes: signal generation, signal amplification, and signal radiation through an antenna. Each is considered in more detail below.

5.2.1 Transmitter Signal Generation

In modern radio transmitters, this function is typically performed by a digital signal processor because such devices allow for the optimization of the transmitted signal. This section describes the factors that affect interference behavior of transmitter – in terms of improving its recovery under interference,[4] as well as causing interference.

5.2.1.1 Signal Bandwidth

System bandwidth tends to vary with the application and the required throughput. Some channelized systems use channel bonding, a technique to provide a higher throughput by expanding the effective width of the RF channel. For example, the High Throughput mode of IEEE 802.11[5] has an optional mode of 40 MHz transmission which doubles the throughput by bonding two adjacent 20 MHz channels. Such bonding does not improve the spectrum efficiency since the bits per Hz ratio is not significantly affected. In fact, spectrum use may be sub-optimal when the interference from other devices in such broader channels are taken into account.

[3]This limitation on channel capacity clarifies that there is no technical panacea that will deliver unlimited bandwidth from limited spectrum. David Reed's "green is unlimited" is only true at the transmit side. The Shannon-Hartley theorem states that a receiver is limited in its ability to separate green from green.

[4]The ability to withstand interference.

[5]Also known as "802.11n".

This is a major concern in the 2.4 GHz band where the number of available channels is very limited.

An advantage of a narrow signal bandwidth is a lower probability of overlap with an interfering signal, but this is offset to a large extent by the more severe impact of frequency selective fading. The latter factor can be compensated for, to some extent, by receiver diversity. The reverse obtains for wideband signals. The balance between these two factors depends on the actual bandwidth used, the properties of the nearby interferers, and the properties of the propagation channel. If the latter is highly frequency selective, there may be little benefit in using a wideband signal.

The probability of interference from other transmitters increases linearly with bandwidth, but these signals are also subject to frequency selective fading, although the fading parameters are likely to be different. This suggests an advantage for the wideband transmitter: narrowband interferers may suffer loss due to selective fading and thus will affect their victims less. Transmit and receive diversity may well increase this advantage.

Compared to a wideband signal, a narrowband signal has a lower probability of affecting another system. However, this is offset to some degree by the power density difference: for the same total power, the narrow band signal has a higher power density; and therefore, the damage it does is likely to exceed that of a wideband signal.

There are four different combinations of interferer and victim bandwidth and these yield four different cases of sharing effects in the frequency domain (see Table 5.1).

Table 5.1 shows that there is an asymmetry in interference effects between narrowband and wideband systems that cannot be overcome on a purely bi-lateral basis: given comparable power levels, narrowband systems will always be more likely to "damage" wideband systems rather than the other way around. Wideband systems can avoid such damage only by increasing the robustness of their signaling at the expense of data throughput – e.g. by coding and spread spectrum techniques.

The preceding addresses the case of two systems being in interference range. Another factor to be considered is the probability of this being the case. For the same RF power, the power density of a narrow band signal is higher than that of a wideband signal and, therefore, it affects victims at larger distances than a wideband signal. The interference distance increases with the square root (or higher root, depending on the environment) of the RF power. Since the number of potential victims goes up with the square of that distance, the number of potential victims

Table 5.1 Impact of bandwidth differences on interference

		System A (interferer)	
		Narrowband	Wideband
System B (victim)	Narrowband	Proportional impact on victim	Little impact on victim
	Wideband	Large impact on victim[6]	Proportional impact on victim

[6] Within certain limits: if the BW ratio is large and B uses OFDM, it may be able to handle narrowband interference.

increases with a fractional exponent of the increase in power density that goes with a reduced bandwidth. For the same transmission power, the number of potential victims is given by

$$N_{vict} = r_{BW}^{\frac{2}{PLE}} \tag{5.2}$$

in which r_{BW} is the bandwidth reduction factor and PLE stands for the applicable pathloss exponent.

Adding the frequency dimension changes this picture: the probability of interference occurring in the frequency domain goes down with the decreasing bandwidth.

$$N_{vict} = r_{BW}^{\frac{2}{PLE}-1} \tag{5.3}$$

In practice, there may be little difference between the interference potential of wideband and narrowband systems of comparable RF power output.

5.2.1.2 Modulation Schemes

Radio systems use different modulation schemes, depending on the need for throughput, robust operation, power consumption, and cost of implementation, to name the four main design considerations. Some modulation schemes offer more interference resistance than others: in general, higher order modulation schemes have lower interference resistance.

Spectrum sharing benefits from efficient use of frequencies: ideally, the spectrum mask of a transmitted signal should be square; steep sides and a flat top. However, every modulation scheme produces sidebands of some form or other. Even simple ON/OFF Keying (OOK) of a transmitter carrier may generate a wide spectrum that could affect many other systems operating in the vicinity. In general, the width of that spectrum is proportional to the symbol rate and therefore inversely proportional to the symbol duration. Packing more data bits per symbol – as is done in higher order modulation schemes – allows increasing the data rate without increasing the used bandwidth. Minimizing the unwanted emissions contributes not only to spectral efficiency in that it allows more channels to be packed in the same frequency band, it likewise reduces the interference seen by other systems. A well-known example of such a modulation scheme is OFDM: the narrow subcarriers of such a signal have correspondingly narrow sidebands – the latter determine the total sideband energy of the OFDM signal – and thus its interference potential (Fig. 5.3).

Another advantage of OFDM is that the OFDM signal has noise-like properties for receivers that are not "trained" to the channel. Therefore, noise reduction schemes like Maximum Ratio Combining are effective in reducing OFDM interference.

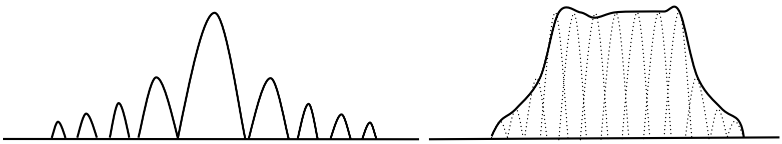

Fig. 5.3 Typical spectra for OOK and OFDM

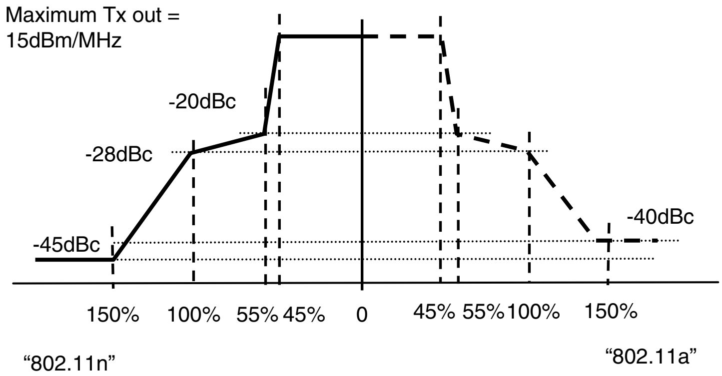

Fig. 5.4 Transmitter output spectrum masks (IEEE 802.11)

5.2.1.3 Transmit Spectrum Filtering

A transmission filter shapes the output spectrum of a transmitter – e.g. by reducing spurious emissions, sidebands, and/or increasing the steepness of the resulting spectrum. Typically, radio regulations determine the mask in which the actual transmission spectrum has to fit. See Fig. 5.4 for an example of such a mask.

A major parameter in spectrum sharing is the spillover of the transmitted power into frequencies outside the intended channel. Such spurious emissions[7] may contain modulation products such as third order Intermodulation products that affect nearby systems. In case of similar systems sharing a frequency band, this is called the Adjacent Channel Leakage Ratio or ACLR. ACLR is a function of frequency; it is measured in dBs for a given distance to the edge of the transmitter's channel.

Transmission filters reduce the width of the filtered signal. Such a reduction can provide significant reduction in interference seen by other systems – this is notably true for the sidebands: here a reduction pays off in terms of reduced adjacent channel interference to other systems. However, such filtering may change the information content of the signal: as redundancy is removed, new, unwanted, information

[7] ITU-R Recommendation M.329–7 defines spurious emissions as "Emission on a frequency, or frequencies, which are outside the necessary bandwidth and the level of which may be reduced without affecting the corresponding transmission of information. Spurious emissions include **harmonic** emissions, **parasitic** emissions, **intermodulation** products and **frequency conversion** products."

may be added that effectively reduces the power of the error-free signal. This may reduce the ability of a receiver to correctly receive such a shaped signal.

As with many other factors discussed here, reducing one effect may cause the increase of another effect. Therefore, the specification of transmission filters is always a compromise between the need to reduce unwanted output and to maintain wanted output components.

5.2.1.4 Pre-distortion to Adjust for Amplifier Behavior

As noted above, transmitter signal generation may lead to unwanted effects on the output spectrum. Similarly, the RF amplifier may introduce unwanted effects – mostly because of non-linearity which, in turn, is the result of power efficiency measures like Class B of C operation.

Pre-distortion of the amplifier input signal may be used to counteract the effect of amplifier – and filtering – induced distortions. By correcting for degradation of the transmitted signal in the RF amplifiers and filters, pre-distortion may contribute to interference resistance of the transmitted signal.

5.2.1.5 Pre-distortion to Adjust for the Propagation Channel

Just as pre-distortion can be applied to counteract known deficiencies of the transmission filters and amplifiers, it may be used to counteract the known or expected deficiencies of the transmission channel – or in the case of MI(M)O systems, the deficiencies of the paths between transmitter and receiver.

By correcting for degradation of the transmitted signal in the channel or transmission path, pre-distortion contributes to interference resistance of the transmitted signal. The degree to which this effect obtains depends on the spatial properties of the transmission path: much spatial diversity assures that all paths are independent. This being the case, it is highly unlikely that the benefits of pre-distortion – as seen by a given receiver – will accidentally be realized by an interfering transmitter.

5.2.1.6 Multiple Transmit Ports (Antennas)[8]

Advances in channel modeling and signal processing have opened up ways to use channel diversity to carry multiple signals at the same frequency. This effect can be used to increase the robustness of the information being carried or to increase the signaling rate and so reduce the time needed to transport the information. Physical variety in the space between a transmitter and a receiver means that there is a number of paths between them that may be virtually independent. With separate

[8] See Sect. 5.4.2 *MIMO*.

antennas, these paths together form a "multiple input/multiple output channel" that supports multiple data streams between two points. This allows for the multiplication of a data rate achievable. There are other ways of using multiple input/output channels: they range all the way from one input to n outputs (SIMO) to n inputs and one output (MISO). The same applies to devices: one can think of multiple receivers optimized for receiving the signal from a single transmit antenna and, at the other extreme, one receiver that is optimized for receiving signals from multiple transmit antennas.

In MISO or MIMO systems, multiple input channels may be used to increase interference resistance of the signal, not only by adding redundancy but also by leveraging spatial diversity of the transmission channel between transmitter and receiver: different spatial paths may have different properties that facilitate or impede reception. Multiple input transmission has no impact on a transmitter's performance if the total power summed over all antenna ports is the same as the output power of a single output transmission.

5.2.1.7 Transmission Beam Forming

With beam-forming, the transmitter utilizes the information of the MIMO channel to generate a spatial mapping matrix to improve reception of a signal. A well known example is the IEEE 802.11 High Throughput option – it adopts beam-forming as a technique to provide a higher throughput between a transmitter and one or more receivers.

There are basically two types of beam forming: implicit beam forming and explicit beam forming. Explicit beam forming requires a transmitter to receive the channel matrices or the mapping matrices measured and sent from the intended receiver.[9] Implicit beam forming does not require the receiver to send the transmitter the MIMO channel information. Instead, it uses the fact that the channel between the transmitter and the receiver is the transpose of the channel between receiver and the transmitter. Since the actual channel includes both the transmit chain in the receiver and the transmitter, a calibration procedure is necessary to correct the difference in the measured channel. Beam forming – like the pre-distortion described in the section above – has the advantage of optimal adaptation of the signal to the transmission path between transmitter and receiver.

By maximizing the delivery of information carrying energy to the intended receiver, transmit beam forming contributes to interference resistance of the transmitted signal. Here too, the degree to which this effect obtains depends on the spatial properties of the transmission path: much spatial diversity assures that all paths are independent. This being the case, it is highly unlikely that the benefits of beam forming – as seen by a given receiver – will accidentally be realized by an interfering transmitter. Transmit beam forming may be used to complement the delivery of the wanted signal at a given receiver by the reduction of interfering energy delivered

[9] When the matrices are compressed, it is called the compressed beaming forming.

to other receivers in known directions. The actual effect depends on many factors, including the degree of gain towards the wanted receiver and the number of receivers to be "protected" from interference.[10]

5.2.1.8 Space Time Block Coding (STBC)

Space-Time Block Coding (STBC)[11] transmits multiple copies of a data stream across multiple antennas. On the receiving side, STBC combines all the received copies of signals optimally. One example of STBC is the Alamouti scheme,[12] which works with at least a two-by-one MIMO system. The Alamouti scheme has the benefit of enabling STBC without knowing the channel information to the receiver. STBC provides for a higher SNIR at the receiver, including a non-MIMO receiver and, therefore, this technique offers improved interference resistance. The same is true for Maximum Ratio Combining (MRC) — provided the interfering signal is noise (AWGN) like.

5.2.2 *Transmitter Amplifier: Power and SNIR*

The RF power of transmitted radio signals is subject to many constraints – technical as well as regulatory. Regulatory authorities typically impose restrictions over the radiated power and/or power density, the power sent by the antenna at its direction of maximum gain. The radiated power is related to the transmission power as follows:

$$radiated\ power = transmitter\ power - system\ loss + antenna\ gain$$

The successful reception of a signal requires that:

(a) the received signal strength is above a certain threshold determined by the receiver sensitivity – this level is an absolute level that is expressed in dBm.
(b) the received signal strength must exceed the noise or interference level by a certain margin that is specific for a given modulation scheme or transmission rate. The latter is called the Signal to Noise plus Interference Ratio (SNIR). It is a relative value expressed in dB.

In general terms, interference may be considered to raise the noise floor of a system. Devices may transmit at a higher transmission power to improve the signal-to-noise ratio (SNR) or signal-to-interference ratio (SIR) as in requirement (b). Therefore, transmission at a higher power level decreases the bit error rate, which in turn improves the channel utilization efficiency. However, a device which increases

[10] See Perahia and Stacey [101].
[11] Ref. [101].
[12] See Alamouti [11].

its transmission power for a higher data rate or a lower error rate increases the interference level for other devices. This is offset to some degree by the reduced time needed to get the same amount of data delivered to the receiver.

However, because the reduction in throughput of the victim systems varies inversely with the throughput gain[13] of the interferer, there is no net gain if interferer and victim belong to the same type of device and to the same network. Increasing or reducing transmission power of all devices in a network has little impact on the overall co-channel interference, because signal and interference change by the same amount: the SIR remains the same.

5.2.3 *Transmission Antenna*

5.2.3.1 Antenna Directivity

Antennas transmit and receive signals over the medium. The size of antenna depends on the carrier wave length. Antennas do not add power to signals. Instead, they change the signal strength in various directions, adding more power in some directions and reducing power in other directions. The resulting spatial distribution of the signal strength gives a characteristic antenna pattern.

The solid angle in which an antenna concentrates the radiated energy is given by

$$\Theta = \lambda / D \ (radians) \tag{5.4}$$

in which λ is the wavelength of the signal and D is the antenna size.

Antenna gain G is typically expressed in decibels (dBi) relative to an isotropic antenna:

$$G = 10 log \frac{2}{(1 - cos\theta)} \tag{5.5}$$

Antenna patterns vary with the form of the antennas. When standing vertically, a dipole antenna has an antenna pattern of 360° in the horizontal plane. Vertically, its beam-width is about 75° – measured the – 3 dB points in the vertical plane. In the horizontal plane, a dipole antenna has a gain of 2.14 dB over isotropic antennas.

Antennas can be classified as omni-directional and directional. Omni-directional antennas provide a 360° horizontal antenna pattern, while directional antennas can strengthen RF power in particular horizontal and vertical directions.[14] Omni-directional

[13] For a given bandwidth, the relationship between throughput and SNIR is logarithmic and therefore this applies only to a homogeneous population of devices. In a heterogenous population, those using higher modulation schemes will suffer more than the interferer gains.

[14] An example is the pencil beam of a tracking radar.

antennas are used in scenarios where coverage is required in all directions in horizontal planes from the antennas. An omni-directional antenna is easy to install and easy to attach to devices. An omni-directional antenna is usually vertically polarized. It would be impossible to apply cross polarization to combat interferences. An omni–directional antenna with a low gain covers more area near the device and is more likely to successfully receive a signal especially in an indoor multi-path environment.

Directional antennas concentrate the RF power in a particular direction which allows larger distances between transmitter and receiver – at the expense of azimuthal coverage. A directional antenna is mostly helpful in Line of Sight (LoS) or near line-of-sight (NLOS) conditions, which, in an indoor environment, includes hallways and corridors. Figure 5.5 gives an example of a radar (=high gain) antenna pattern.

Directional antennas can be very good tools to manage interference: the tighter the transmitted beam they produce, the smaller the chance that this beam will intersect the location of the receive antenna of another system. An example will make this clear (see Fig. 5.6). Consider a transmitter with a given output power radiating into free space and potential victim systems with an interference threshold of x.

In the case of the omni-directional antenna, the transmitter's signal strength drops to x at distance R. The circular area enclosed by the unit radius R is πR^2.

In the case of the directional antenna with an aperture angle of φ, the distance R_φ at which the transmitter's signal strength drops to x is proportional to the antenna

Fig. 5.5 Example radiation rattern -high gain radar antenna (plan view, not to scale)

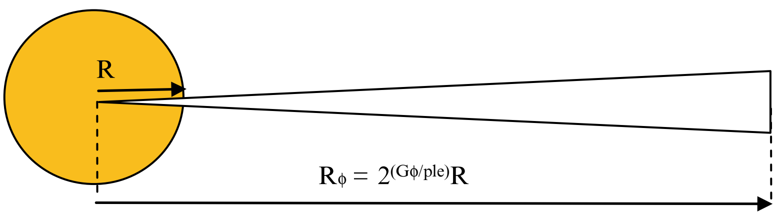

Fig. 5.6 Interference areas for an omni-directional and a directional antenna for the same power output level

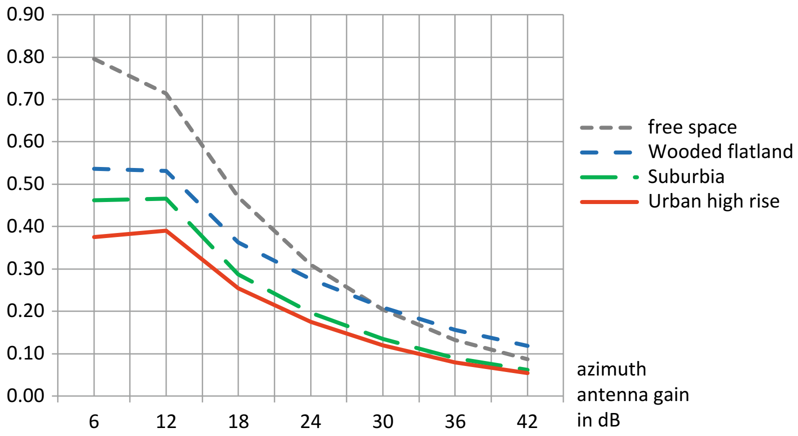

Fig. 5.7 Relative antenna "footprint area" vs. antenna gain for different propagation environments[15]

gain: $R_\varphi = R\log_2(G_\varphi)$ in linear terms or $R_\varphi = R^{(G\varphi/6)}$ in dB.[16] The area A_ϕ enclosed by the triangle is given by:

$$A^\varphi = \frac{(\alpha \times \pi R^2)}{360} \tag{5.6}$$

Operating range R_φ grows quickly with increasing antenna gain, and this suggests an ever increasing interference range and therefore an increasing interference area. However, propagation and horizon effects prevent this. Figure 5.7 shows how the ratio of the footprint of a directional antenna varies with the antenna gain and the pathloss exponent for different environments. Both complex pathloss and the horizon effect have been taken into consideration by defining multiple breakpoints and different pathloss exponents for free space, open landscape, suburban and urban environments. For higher pathloss exponents, the interference distance drops to very low values and the above ratio becomes even more favorable.

However, there is a catch: the improved sharing behavior is only achieved if the antenna gain corresponds to directional selectivity. Since most victim devices may be assumed to lie more or less in the horizontal plane, the directional selectivity has to be in the azimuth. The elevation aspect is, in practice, largely irrelevant.[17] Thus, it is the azimuthal gain that has to be used in deriving the interference footprint ratios.

[15]The antenna angle varies with the gain. The figure is based on: 6 dBi = 72°, 12dBi = 36°, 18dBi = 18°, 24dBi = 9°, 30dBi = 4.5°,36dBi = 2.2°, 42dBi = 1.1°.

[16]These numbers apply for free space conditions only.

[17]A clear exception is the high power weather radar: changing the antenna elevation greatly affects the ground level projection of its pencil beam at a given antenna gain.

Antenna directivity contributes to interference resistance in two ways: by improving system gain and therefore the SNIR at the intended receiver and, secondly, by reducing the population of potential interferers: narrower antenna angle means it will see less neighboring systems. Because the antenna footprint area actually decreases with longer distance and therefore with higher pathloss exponents, the high gain antennas improve a system's interference resistance more than low gain antennas.

5.2.3.2 Antenna Polarization

Antennas determine the polarization of the signal they transmit (or receive). Depending on the design of the antenna, the polarization may be anything between linear and circular. The polarization is specific for a given antenna – two or more antennas with different polarizations may be combined into a single array that supports multiple polarizations. Theoretically, the polarization loss between a perfect plane wave and the antenna is given by:

$$F = 1 - \cos^2(\tau) \tag{5.7}$$

where τ is the angle between the wave plane and the antenna plane F varies from 0 for 0° (full coupling) to 1 at 90° (no coupling).

For non-perfect plane waves, the loss factor is a complex function of the two wave angles relative to the antenna orientation. In practice, perfect plane waves will not occur: reflection of a plane wave at angles other than 90° will change the plane polarization into a somewhat elliptical polarization. This loss of polarization increases with the number of reflections. Even small disturbances in the RF medium will affect the polarization of the signal. Therefore, coupling loss will vary, e.g. between 30 dB for a highly polarized signal to as little a 3 dB for a circular or randomly polarized signal. This practical limit to polarization coupling and de-coupling has consequences for interference between RF systems.

In a (near) reflection free and scatter free propagation path transmitter and receiver, polarization is an effective means of increasing interference resistance. The effectiveness improves when other propagation paths are either blocked or subject to much scattering. Plane polarization may reduce the impact of interference, notably when the interfering wave is highly polarized. In practice, this will not occur often – if at all.

5.3 Signal Propagation

This section considers only the factors of channel propagation that impact interference resistance or causation.

5.3.1 Propagation Path Loss

Propagation loss in a communication channel is a major obstacle for reliable wireless communications. Propagation losses vary significantly with time and space conditions, which make it a factor that is nearly completely beyond the control of communication system designers. Four phenomena affect the propagation of an RF signal: attenuation, reflection, diffraction, and scattering. The impact of each of these four is proportional to the frequency of the propagating signal.

The attenuation of an RF signal in space is a direct consequence of its three-dimensional (usually spherical = omni-directional) propagation of the signal from the transmitting antenna. The field strength of the signal is inversely proportional to the size of the sphere or sphere segment with a radius that is the distance to the transmitting antenna. This applies only in free space. When the propagation space is littered with objects that absorb, block, reflect or scatter the signal, the effect is a decrease in the received signal level that exceeds the simple square law. Instead, the pathloss exponent will be three or far more, depending on the environment. This applies notably in the case of diffraction limited propagation (see Sect. 5.2.2).

Signal attenuation affects all radio signals and therefore it does little to affect the interference resistance of a system. However, because of the higher pathloss exponents observed at larger distances, nearby interferers typically will cause more interference than the same interferers at larger distances. Because of the higher pathloss exponents observed at larger distances, nearby victims will usually experience much more interference than similar victims at larger distances. Although this is true as a rule of thumb, wavelength and atmospheric effects may have a major impact: ducting caused by inversion layers reduce the pathloss exponent to a degree that depends on the wavelength of the RF signal, as well as the refractive index that obtains at the layer boundaries.

5.3.2 Signal Degradation by Reflections, Diffraction and Scattering

RF signals reflect on the obstacles they encounter during propagation. When a radio frequency signal meets a large object during propagation, the signal changes its propagation direction and reflects from the object. Reflection occurs if the surface does not absorb RF energy very well and the wavelength of the signal is considerably smaller in size than the object, e.g. as in the case of buildings or mountains.

Diffraction is the bending of waves around the corners when a signal in propagation encounters an object both large and sharp-edged in terms of the propagating wavelength. Secondary electromagnetic waves around the object are generated, whose directions differ from the incoming signal. If an obstacle exists in the path between the transmitter and the receiver, some energy can pass through because of

diffraction on the top edge of the obstacle. Diffraction is one of the reasons why, even behind a reflecting object, some signal strength will be observed: typically in the order of 20 dB less than the incoming signal.

Scattering takes place when a signal propagates in a medium with many small objects such as foliage and rain drops whose sizes are much smaller than the wave length of the signal. The resulting signal will then deviate from its original direction. Under favourable conditions, scattering may reduce signal strength by as little as 10 dB in directions away from the line of sight between transmitter and receiver.[18]

Signal degradation affects all radio signals and, therefore, they do little to affect the interference resistance of a system. Although the environment itself does not contribute to interference, propagation effects may cause local enhancements of interfering signals at a given receiver.

5.3.3 Frequency Selective Fading and Delay Spread

Because of the frequency dependency of propagation impairments, the actual signal at the receiver will vary in space and in time – the latter being caused by motion of the transmitter, the receiver or the objects in the propagation path. This variation is called fading.

Small-scale or Rayleigh fading refers to path loss variations over fractions of a wavelength or several wavelengths. Large-scale or Rician fading caused by e.g. shadowing, refers to pathloss variations on a scale corresponding to many wavelengths.

A receiver receives both the direct, primary signals, and the secondary ones, those reflected, diffracted or scattered. The combination of a primary signal and the secondary ones is known as a multipath signal. Depending on the phases of these signals, multipath effects can result in increase or decrease of the combined amplitudes of the received signals. Since the distance travelled by the various signals is different, there will be a time difference in their reception. This is called the delay spread; typically, this is measured in tens of nanoseconds (for indoor environments) or microseconds (for outdoor environments). A 100 ns delay spread corresponds to a path difference of 30 m. The amount of delay spread varies with the amount of obstacles or infrastructure between the transmitter and the receiver. For example, the delay spread is larger for a manufacturing floor with a lot of metallic structures than that in a home environment.

Lower bit rates tolerate longer delays. For example, the limit of the delay spread for a frame error rate less than 1% at 11 Mbps is about ns, 225 ns for 5.5 Mbps, 400 ns for 2 Mbps and 500 ns for 1 Mbps. Some receiver designs, notably for single carrier systems, use an equalizer to counteract these delay spread faults.

[18] See de Jong [73].

Receivers for multi-carrier systems – e.g. using OFDM – rely on long symbol duration to reduce delay spread effects.

Frequency selective fading affects all radio signals and, therefore, it does little to affect the interference resistance of a system. However, specific adaptations of the transmitter signal – e.g. the use of an equalizer – and other specific measures in the receiver may considerably impact the interference resistance of a particular system. This is true in particular for equalizer based receivers, see Sect. 5.4 for more details. Directional antennas can be used to minimize the effect of delay spread by limiting the number of reflecting paths that affect the receiver. Circular wave polarization is another measure to counteract the delay spread by cancelling the first reflections, since the circular direction of the reflected signal is opposite to the direct one.

Although fading and delay spread on their own do not contribute to interference, propagation effects may cause local enhancements of interfering signals at a given receiver. However, specific adaptations of the transmitter signal and specific measures in the receiver may considerably affect the interference resistance of a particular system. In general, the more specific the adaptation, the greater its positive impact on the system's resistance to noise-like interference.

5.4 Signal Reception and Interference

Successful signal reception requires three conditions to be satisfied: the received signal must be detectable by the receiver and exceed the noise in the receiver itself, the received signal must exceed the noise in the propagation channel and, thirdly, the received signal has to exceed the interference present in the propagation channel. In addition, receivers have the property of selectivity: the ability to separate one signal from another signal. This plays a major role in receiver performance under conditions of interference.

5.4.1 Noise

5.4.1.1 Noise Figure and Noise Floor

While noise can emanate from many sources, when looking purely at the receiver, the noise is dependent upon a number of elements. One element is the inherent thermal noise figure generated by the receiver circuits. This noise figure – N_{TH} – is the difference between the noise at the input of the receiver and the noise at its output. This minimum noise contributed by the receiver itself is a function of its bandwidth: At room temperature (~2,900 K), the receiver noise floor is given by

$$NF_{290} = -174 + N_{TH} 10logB \tag{5.8}$$

where NF is the figure of the receiver and B is the bandwidth of the receiver in Hz.

A low noise figure allows the receiver to operate with lower input signals, but it also increases its sensitivity to interference. A high noise figure requires a higher input signal, but it also makes a system less sensitive to interference. Since the noise floor varies with the receiver's bandwidth, the impact of interference is determined by the ratio between the width of the interferer's spectrum and the receiver bandwidth.

5.4.1.2 Receiver Signal to Noise Ratio

In addition to receiver sensitivity, a receiver needs a minimum signal to noise ratio to achieve a certain bit rate. Signal to noise ratio is the difference between the desired received signal strength and the noise floor:

$$SNR_{RX} = 10 \log\left(\frac{RSS}{noise\ floor}\right) dB \tag{5.9}$$

In which RSS stands for received signal strength. Receiver design plays a major role in the determination of the inherent SNR of a receiver: better components, higher resolution digital conversion and processing, all contribute to a lower = better receiver SNR.

A better receiver SNR means that the receiver will be better able to separate its own noise contribution and the received signal. The latter may include channel noise or noise-like interference.

5.4.2 Receiver Sensitivity

Receiver sensitivity is the threshold for the minimum received signal power, for a given data rate to be received with a given error rate or better. The latter equates to a certain SNR at the output of the receiver. A higher data rate requires a higher sensitivity and vice versa. For a given signal strength, a higher (= better) receive sensitivity means a higher frame success rate. For a given transmitted power limit, the higher receiver sensitivity gives a better operating range.

Receiver sensitivity is defined in the absence of specific interference signals – but it does take into account the noise level of the system – i.e. it includes the receiver's thermal noise, the channel noise, and the noise contributed by noise-like interference sources.

Receiver sensitivity is determined by the above noise factors, as well as by the gain of the receiver circuits: more gain means more signal, but also more noise. Therefore, receiver gain must be tuned to the noise level of the RF channel to achieve optimum results. In digital receivers, the automatic gain control (AGC) of the pre-amplifiers that feed the digital to analog (D/A) conversion circuits determines only part of the overall gain – signal processing applies further gain adjustments. In general, receiver sensitivity is given by:

$$S_i = 10\log\left(k_{(T\alpha+Trx)} \times B \times SNR\right) dB \tag{5.10}$$

where k = Bolzmann constant, T_a = equivalent noise temperature in [K] of the source (e.g. antenna) at the input of the receiver, T_{rx} = equivalent noise temperature in [K] of the receiver referenced to the input of the receiver, B = bandwidth in Hz and SNR is the required SNR at the output of the receiver.

Thus, $\mathbf{S_i}$ varies with the desired data rate. Receiver sensitivity affects interference resistance through the receiver gain factor: more gain means less interference resistance. It follows that higher data rates, which require higher receiver sensitivity also suffer from reduced interference resistance.

5.4.3 Co-Channel Interference

Co-channel interference arises from transmissions by devices operating in the same channel as the receiver. It is typically the most significant type of interference a receiver is exposed to. The interfering signals may fully or partially overlap the receiver's channel. In addition, each interference source has its own unique characteristics, so that its impact varies; this is also true of the techniques to combat the interference. Depending on the receiver and the interfering signal, interference may be treated as a source of additional noise – essentially raising the level of the noise floor. The ratio of signal to noise + interference is called the SNIR. Like the minimum useful sensitivity, the required SNIR varies with the signal rate.

Co-channel interference, even when it is relatively weak, directly affects the SNIR of a receiver and therefore its ability to perform its intended function adequately; it is the main limiting factor of spectrum capacity. The interference impact can be estimated using the Shannon-Hartley theorem which states that, for a given bandwidth, the capacity of a channel is proportional to the logarithm of the SNR – see Sect. 5.1. More transmitters in the vicinity of a receiver means more interference and therefore a lower SNIR and a lower data rate. Spectrum capacity is a local phenomenon. As seen by a given user, it is a function of the number of devices transmitting on the same channel, their power levels and their distance.

Addressing co-channel interference, therefore, requires isolation of the victim and the interfering source(s); this can be achieved in a variety of ways, e.g. by frequency

planning,[19] by increasing the SIR of the victim system, by coding the wanted signal as in CDMA systems or by separation in time as in TDMA or CSMA systems.

5.4.4 Receiver Selectivity

Receivers have the property of selectivity: the ability to separate one signal at the desired frequency from another signal at another frequency. Receiver selectivity has a number of elements.

5.4.4.1 Receiver Blocking

Receiver blocking is an effect caused by a strong out-of-band signal, present at the input of the receiver. It reduces the receiver's ability to detect an in-band wanted signal. The blocking signal reduces the specified receiver sensitivity by a certain number of dB's. Receivers with a large blocking protection ratio have good interference rejection properties.

5.4.4.2 Intermodulation Sensitivity

Intermodulation products affect the ability of a receiver to maintain its bit error rate (BER) in the presence of other signals. Intermodulation products occur when out-of-band interfering signals are mixed with in-band signals because of non-linear behavior of the receiver front-end. Notably the 3rd order Intermodulation products are important because their amplitude at the receiver (front-end) output rises as the third power of the unwanted input signals. The input level at which the third order products begin to affect the receiver output is known as the Intermodulation Distortion threshold or IMD threshold. The difference to the minimum useful signal level is known as the IMD dynamic range. Interfering signals may give rise to IMD. Receivers with a large IMD dynamic range have good interference rejection properties.

5.4.4.3 Spurious Response

Spurious response in a receiver occurs when unwanted signals, having frequencies other than the tuned frequency, produce a receiver output as if they were wanted signals due to mixing with harmonics of the receiver's oscillator. Inducing spurious response behavior typically requires very strong unwanted signals such as occur in electronic warfare. It usually is of no concern in spectrum sharing analysis.

---

[19] See Chap. 10, Radio Resource Management.

5.4.4.4 Adjacent Channel Rejection

Adjacent Channel Rejection is a measure of the receiver's ability to reject signals from adjacent channels, while receiving a wanted signal on its tuned frequency while maintaining its desired BER. Sometimes this selectivity is referred to as Adjacent Channel Rejection ratio or ACR ratio, which is useful in assessing the capacity of a multi-channel system. Receiver selectivity is largely determined by the receiver's input filters: better filters mean steeper filter slopes and therefore a better rejection of adjacent and further adjacent channel signals.

Adjacent Channel signals may consist of or include interfering signals. Therefore, a good ACR improves the interference resistance of a receiver.

5.4.5 Inter-Channel Interference and Partial-Band Interference

Inter-channel interference may occur because the spectrum mask of the transmitter of the interfering device(s) does not completely suppress the power outside its channel bandwidth. Such interference, however, is usually far less significant than co-channel interference because the suppression of the leaking power often results in an interference of low magnitude. Some overlap in channel use is acceptable and the performance of a system in terms of throughput for a given total bandwidth improves until some optimum overlap value.[20,21]

Partial-band interference[22] occurs when the interfering signals overlap partially with the channel of the receiver. One of the significant characteristics of such interference is that it is non-white and may be correlated with the wanted signal. Therefore, it may have severe impact on the reception of the wanted signals.

As discussed above, inter-channel interference and notably partial-band interference may have effects that exceed expectations based on simple signal strength. Since there is little a receiver designer can do against this type of interference, allowing dissimilar systems to share a frequency band reduces the efficiency of spectrum utilization.

5.4.6 Receiver Spatial Diversity and MRC

In general, spatial diversity can be used to improve the SNR at the receiver. The conventional "receiver diversity" approach makes use of the property of channel orthogonality that is more fully exploited in MIMO systems[23]: antenna signals are

[20] See Teotia [121].

[21] This observation is the basis for channel assignments that use the frequencies in the gaps between two fully separated adjacent channels.

[22] See Betz [26].

[23] See also Sect. 5.4.2 *MIMO*.

fed into separate receiver paths for evaluation of the signal strength. The stronger signal is used for further processing because the received signal power tends to correlate with the actual strength of the received signal.

Maximal Ratio Combining (MRC) maximizes the quality of receiver's output by using a weighted sum of the signals received on all antennas. The weight for each signal is different and it is chosen to be proportional to the root-mean-square of the signal strength and inversely proportional to the mean square of the noise level. MRC leads to a combined signal that is stronger than the one received with maximum power. This improvement in SNR results in a better signal reception range and therefore leads to better coverage.[24] However, interference affects the efficacy of MRC. A noise like interfering signal may reduce the MRC gain and therefore the obtainable SNR improvement.

5.5 Advanced Transmission Techniques

Recent advances in signal processing theory and practice have led to the emergence of two new concepts that are both based on "parallelism": OFDM, which uses parallelism in the frequency domain, and MIMO, which uses parallelism in the spatial domain. MIMO – "multiple input – multiple output" – exploits the fact that, in a complex transmission environment, there exist a large number of approximately orthogonal propagation paths between two locations. By employing multiple antennas that are spaced appropriately, these orthogonal paths can be exploited to provide more capacity or redundancy – or a mix of both.

OFDM and MIMO have given rise to major improvements in the capacity of RF systems. Both systems require that transmitter and receiver are designed to operate together. Therefore, it makes good sense to talk about OFDM and MIMO systems rather than ditto receivers.

5.5.1 OFDM

OFDM systems spread energy and information over a number of simultaneously transmitted subcarriers. Each subcarrier transports only a fraction of the total information – i.e. a small number of bits per symbol – and this allows the OFDM symbols to be relatively long in duration. This improves the tolerance for delay spread and other channel impairments. Coding across the subcarrier improves overall robustness. "Water filling" in which subcarriers are given power levels that

[24] See Mallik [90].

compensate for frequency selective channel impairments allows throughput to be optimized. Taken to extremes, this means that subcarriers, which are not productive due to heavy fading or interference, may not receive any power at all.

In an interference-rich environment with a mix of interfering sources of different properties, OFDM systems may have an advantage because they are much better able to withstand narrowband interference than e.g. equalizer-based receivers.

5.5.2 MIMO

In a MIMO channel, each input and output corresponds to an antenna that couples a device to the channel. MIMO leverages the multiple, independent paths which exist between two locations within a spatially diverse environment.[25] MIMO can be leveraged to increase channel capacity or transmission robustness – or some mixture of both.

MIMO systems have been under development for many years, since its inception at Bell Laboratories in 1984 by Jack Winters. MIMO technology has been successfully deployed in third generation cellular systems to improve both downlink and uplink throughput. Recent ratification of IEEE 802.11n specification and the resulting popularity of IEEE802.11 High Throughput Access Points and Clients have greatly increased the user base of MIMO technologies in commodity wireless systems.

A MIMO device uses multiple transmitter and receiver antennas to allow for multiple simultaneous data streams. A multiple input device simultaneously sends more than one RF signal into its transmitting antennas. A multiple output device simultaneously receives more than one RF signal from its receive antennas. The technology provides improvement in overall system performance by increasing data throughput via spatial multiplexing, increasing range via spatial diversity and increasing signaling robustness by adding redundancy. In addition, MIMO systems may employ beam forming: a MIMO transmitter can drive its multiple antennas to direct the signal power towards the target receiver with a phase-shifting algorithm, which in theory can increase the effective signal power by the square of the number of antennas. Such increase is the combined result from both the power gain of multiple transmissions and the array gain of the beam-forming.

Another MIMO technique is to apply subcarrier-based Maximal Ratio Combining across all sub-channels. In a MIMO system, MRC selects signal components from multiple antennas based on signal strength at each frequency. MRC can improve the overall gains significantly, which is especially applicable in a multipath environment. One of the benefits of this approach is that it does not require the transmitter to be a MI(MO) device. However, transmitting the same data over multiple antennas does not provide the benefit of spatial multiplexing.

[25]This means that MIMO does not work in free space and in line of sight conditions.

5.6 Ultra-Wide-Band Transmission

Ultra-Wide-Band transmission promised gigabit data rates at high electrical efficiency. Notwithstanding early support from the FCC in the form of flexible rule-making, it has never seen the rate of adoption its early proponents foresaw. Nonetheless, it is worth looking into the spectrum sharing between wireless LANs and Ultra-Wide-Band systems – if only because the latter present an extreme approach to spectrum sharing: very low power at very large bandwidths. In effect, there is no dynamic sharing mechanism involved. The same is not true for frequencies above 3.1 GHz. Here Ultra-Wide-Band devices have to implement a "detect and avoid" (DAA) mechanism to protect radar and broadband wireless access systems.[26] This mechanism is a bit like the DFS mechanism used by wireless LANs to protect radars in the 5 GHz band.

Ultra-Wide-Band communications technology has a long history that started in the 1960s with contributions from academia as well as military research organizations. US patent 3,728,632 dated 17th April, 1973, is considered a landmark patent in Ultra-Wide-Band communications. An excellent historical overview and technical description was given by Terrence Barrett of UCI in 2000[27] – at the time the FCC started its enquiry into developing rulemaking for this technology. The FCC's motivation was that Ultra-Wide-Band promised more efficient use of ("scarce") spectrum resources.

Using Ultra-Wide-Band for communications purposes means making a fundamental choice in favor of bandwidth rather than signal to noise ratio. The Shannon-Hartley theorem says that channel capacity is determined by the product of the bandwidth and the logarithm of the signal to noise ratio. Whereas conventional communications systems climb the exponential slope of the SNR factor,[28] Ultra-Wide-Band systems expand in the linear domain of the bandwidth factor. In doing so, Ultra-Wide-Band systems, unlike conventional systems, necessarily overlap in the frequency domain with many other systems – licensed as well as license exempt. That latter fact lies behind the reason for the massive political resistance which Ultra-Wide-Band generated once the FCC and other regulators began to address the matter in earnest.

In 2002, the FCC released its first Report & Order[29] on the matter which defined Ultra-Wide-Band technology[30] as having a fractional bandwidth[31] of at least 25% of the operating frequency and a maximum bandwidth of 1.5 GHz. The latter restriction takes effect only above 6 GHz. The Report & Order also defined a number of

---

[26] ECC Report 120 [2].

[27] See http://www.ntia.doc.gov/osmhome/uwbtestplan/barret_history_(piersw-figs).pdf.

[28] Because of the exponential SNR function, the power needed to increase the channel capacity also rises exponentially. Ultra-Wide-Band avoids that and delivers a large bandwidth at power levels that allowed integration with semiconductor logic.

[29] See FCC 02–48 – Report&Order on UltraWideBand Transmission Systems.

[30] See FCC 47CFR15.517 and following sections.

[31] The fractional bandwidth is defined as 2(fH-fL)/(fH+fL) at the -10 dB points of the spectrum envelope.

uses for Ultra-Wide-Band technology, including "Communications & Measurement systems" which allowed devices to operate in the frequency band 3.1–10.6 GHz. The equipment had to be designed "to ensure that operation can only occur indoors or it must consist of hand held devices […]".

At the beginning of its long struggle from promise to reality, Ultra-Wide-Band technology was heralded as a radical solution to both spectrum scarcity and the power demands of conventional modulation radio systems. Today, it is another useful RF technology in the armory of the wireless industry. The reason for this situation is threefold: the restrictions on Ultra-Wide-Band operating parameters, necessary to protect established spectrum users; the rapid advance of other radio technologies, such as broadband OFDM; and last but not least, the rapid reduction in the size and power consumption of semiconductor components.

However, Ultra-Wide-Band technology has not disappeared; there are chips and other components on the market that find application for various purposes and in different environments – including industrial communications systems that leverage its interference resistance. Ultra-Wide-Band has been adopted by the USB community.[32]

Ultra-Wide-Band technology is also in use as ground/wall penetrating radar; that type application is not considered here.

5.7 Summary

This chapter makes it clear that there are many physical factors and techniques that affect the transfer of information over the RF medium. See Table 5.2 which summarizes the various factors and techniques discussed in this chapter. The items in bold text generally contribute positively to spectrum sharing. This table suggests the following conclusions:

(a) The variety of these factors and techniques and the variety of their effects greatly complicate the problem of spectrum sharing. Minimizing the differences between sharing systems promises not only to simplify matters but also to improve sharing in practice.
(b) Largely because of propagation factors, there may be a large difference between the interference potential and interference impact. This has to be taken into account in sharing assessments.

If the same spectrum is used by different systems, their interactions are complex and the effects sometimes counter intuitive. Therefore, specific rules may well advantage some systems or designs and disadvantage others. This undesirable effect can be avoided if such rules are limited to total RF power output and, optionally, a cap on RF power density. This subject is explored further in Chap. 12, Sect. 12.3. *A Sharing Metric: the Transmission Unit.*

[32] See www.USB.org. It states that the Wireless USB performance is targeted at 480 Mbps at 3 m and 110 Mbps at 10 m.

Table 5.2 Summary of Physical Factors Affecting Spectrum Sharing

Factor	Interference impact	Interference resistance
Transmitter power and SNIR	More power = more interference	More power = more interference resistance
Transmission antenna	Reduces interference potential	Increases interference resistance
Signal propagation	Neutral	Neutral
Propagation path loss	Reduces interference	Reduces interference resistance
Signal degradation by reflections, diffraction and scattering	May reduce interference impact	Degrades interference resistance
Frequency selective fading and delay spread	Neutral	Degrades interference resistance
Narrowband operation	May reduce interference potential but increase impact	Reduces probability of interference *and* increases interference resistance
Wideband operation	May reduce interference potential *and* impact	Increases probability of interference, degrades interference resistance
Noise	Neutral	Degrades interference resistance
Receiver sensitivity	Not applicable	Degrades interference resistance
Co-channel Interference	Not applicable	Degrades interference resistance
Receiver selectivity	Not applicable	Increases interference resistance
Inter-channel interference and partial-band Interference	Not applicable	Degrades interference resistance
Receiver spatial diversity and MRC	Not applicable	May improve interference resistance
OFDM	May reduce interference impact	May improve interference resistance
MIMO	May reduce interference impact	May improve interference resistance for the same data rate

Chapter 6
Medium Access Etiquettes and Protocols

Accommodating multiple access to spectrum resources is fundamental to wireless networks. In many cases, these access mechanisms operate in the time dimension, sometimes in combination with adaptive selection of the operating frequency. Conventional time-based sharing mechanisms such as TDMA do not fit well with consumer usage, which is generally unstructured. Instead, self-organization based on distributed control is preferable for commodity wireless technologies.

Sensing-based medium access has become a widely accepted principle; it is the basis of "Listen-Before-Talk" (LBT) as well as Detect-And-Avoid (DAA) etiquettes. Using Aloha and the IEEE 802.11 wireless LAN standard as examples, this chapter addresses etiquettes and medium access protocols for commodity wireless technologies. The chapter closes with brief discussions of medium access mechanisms described in the IEEE 802.15.1 standard (Bluetooth) and the IEEE 802.16 standard (WiMAX) and compares them with the IEEE 802.11 medium access control approach.

6.1 Simple Etiquettes: Aloha and "Listen-Before-Talk"

Medium access has been a research topic for many years and many designs have been proposed by researchers and developed by the wireless industry. Some designs rely on separation in frequency, some, e.g. frequency hopping designs, combine this with separation in time, whereas others rely on separation in time only. The latter method may rely on a single point of control or it may rely on distributed medium access control. The simplest form of distributed medium access is the Aloha scheme: it relies on randomization to avoid collisions on the medium. Its throughput is given as:

$$S = G \times e^{-2G} \tag{6.1}$$

where G is the normalized traffic load. Maximum throughput is only 0.184 occurring at G = 0.5 as shown in Fig. 6.1.

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_6, © Springer Science+Business Media B.V. 2011

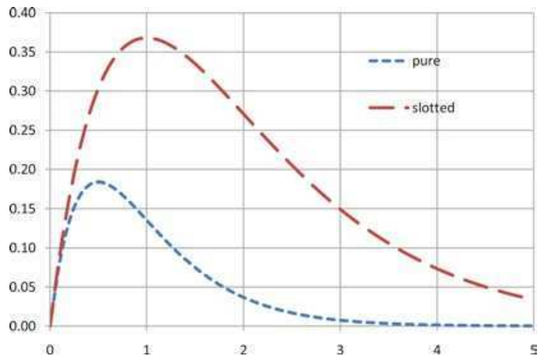

Fig. 6.1 Aloha throughput

Early research showed that a slotted transmission procedure allows a significantly higher overall system throughput than a non-slotted transmission. The throughput of slotted ALOHA[1] is:

$$S_{slotted} = G \times e^{-2G} \tag{6.2}$$

Here, the maximum throughput is 0.368 at G = 1.0. Although the randomization aspect of Aloha proved useful, its throughput was less than desirable, even with slotting applied. The solution proved to lie in the combination of sensing the medium and randomized access to determine the instant of transmission.

A basic method to accommodate multiple access to a shared RF medium is the so-called LBT procedure[2]: before transmission, a device "listens" to the channel to ensure the medium is free. Collisions can still occur if two devices happen to transmit at the same time. Randomizing the instant of medium access, triggered by the medium becoming free, eliminates most potential collisions and it improves throughput to more than 90% of the channel capacity.[3] The actual value depends on the ratio between transmission time and the time needed for medium sensing and the randomized access procedure. The more time is allocated to the randomized access period, the lower the probability of collisions, but the lower the throughput. Thus, designers have to strike a balance between efficiency of channel utilization and successful transmission.

Aloha and LBT are examples on spectrum sharing etiquettes – schemes in which there is no or very little explicit information exchanged between the entities involved. Devices operating under an etiquette do not have to understand each

---

[1] Roberts [109].

[2] Its cousin is Detect And Avoid – it operates in the frequency domain.

[3] Actual throughput depends on the randomization factor as well as on protocol overhead.

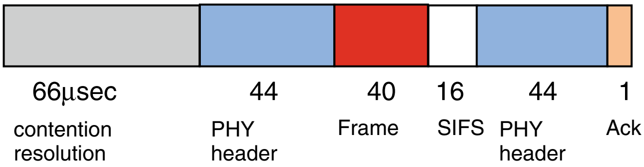

Fig. 6.2 Example of non-aggregated throughput of 802.11n

other's signaling. Etiquettes are therefore simple, but they are also sub-optimal in terms of medium utilization. Optimal medium utilization requires the exchange of information such that participating stations have the same view of the status of the shared medium.

The exchange of this information requires a protocol: with this addition, a spectrum sharing etiquette becomes a medium access control protocol. Even with the addition of medium sensing, randomization cannot avoid all collisions and, therefore, a medium access control protocol has to provide an error recovery mechanism for failed frame transmissions as well. The IEEE 802.11 standard includes this, in addition to the combination of LBT and randomized access, in its medium access method known as Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). In fact, the standard defines both a form of centralized medium access control as well as distributed control; in practice, the latter is dominant.

Spectrally efficient medium access requires overhead – e.g. as delay associated with the random access window and as time needed to transmit acknowledgement and other medium access related information. This overhead can severely lower the bandwidth available for transmitting payload. For example, the top signaling rate of the 802.11n transceiver – 260 Mb/s[4] is almost 5 times that of the 54 Mb/s of an 802.11 g transceiver. Nevertheless, the net throughput of the former is about the same as that of the latter, at least for short and medium-sized frames. The reason is illustrated in Fig. 6.2: the fixed and signaling rate independent medium access overhead is considerable. In this example, 40 µs for transmitting a 1,500 byte data frame at 300 Mb/s requires a total "channel occupancy period" of 211 µs. That is a protocol efficiency of ~20%. Clearly, increasing medium access efficiency plays a key role in realizing the advances in transmission technology. An example is frame aggregation: it is necessary to realize high throughputs made possible with MIMO transmission. This is addressed in more detail below in Chap. 7, Sect. 7.3.

In the following sections, we will first discuss the IEEE 802.11 MAC for wireless LANs as the basis for the remaining sections. That will be followed with discussions on a number of its variations, such as OFDM-A and directional medium access which are designed to address challenges posed by various deployment scenarios.

[4]MCS31 in a 20 MHz channel.

6.2 The IEEE 802.11 Standard[5]

"802.11" is the IEEE standard for wireless LANs operating in the ISM bands at 900 MHz, 2.4 and 5.8 GHz as well as the 5.2 and 5.4 GHz bands.[6] Variations of it are used in the 3,650 MHz band for Broadband Wireless Access, in the Public Safety band at 4.9 GHz and for Intelligent Transportation Systems (ITS) operating in the 5.9 GHz band.[7] The standard defines the lower layers of the protocol stack: the Physical Layer (PHY), which defines the transceiver and its signaling and control features, and the Medium Access Control Layer (MAC), which defines the protocol that governs the exchange of transmission frames, their structure and the content of control and management frames.

The ISM bands present a case of spectrum sharing between dissimilar systems: these bands are used by IEEE 802.11 compatible devices as well as many other types of devices. Therefore, one device's RF signals can become another's interference and cause transmission failures.

Although this was recognized as a potential problem during the development of this standard, the focus was on avoiding collisions of transmissions of the wireless LAN devices themselves. To combat collisions, the IEEE 802.11 MAC adopts the approach of CSMA/CA (Carrier Sense Multiple Access/ Collision Avoidance) with slotted transmissions, as discussed in the preceding section. The slot time of an 802.11 device depends on its operating range as well as hardware properties: it is based on the sum of the propagation delay plus the RX-to-TX turnaround time for the device plus the energy detect time.

There are a number of implicit assumptions that underlie the 802.11 MAC specification. For example, it assumes perfect connectivity between participating devices, which implies omni-directional antennas and the absence of obstacles. These implicit assumptions apply to many wireless LAN deployment scenarios, but they are not always valid. Exceptions usually lead to issues that degrade the performance of a wireless network. Commonly cited examples of these issues include:

- Exposed nodes issue
- Hidden nodes issue
- Near/far nodes issue
- Unfairness and Starvation issues

Some specific issues are fixable with extra measures; RTS/CTS (request to send / clear to send) for the hidden nodes[8] issue for example, while others are not so easily

[5]This chapter is based on the 2007 release of the standard; it incorporates many previous separate Amendments such as 802.11e for Quality of Service and 802.11i for improved security.

[6]Known as the Unlicensed Information Infrastructure bands in the US.

[7]This list is an inclusive generalization; not all bands are available in all countries of the world. Support for other frequency bands – e.g. the TV White Space frequencies in the UHF band–is being developed.

[8]RTS/CTS is spectrally inefficient and it does not work in a static environment – see also Chap. 7, Sect. 7.6.

fixable and remain hot research topics in academia. More details on this aspect of contention-based medium access control are given in Chap. 7, Sect. 7.2.

The IEEE 802.11 standard specifies a contention-based mode of operation using CSMA/CA, which is mandatory, and an optional contention free mode of operation, in which the Access Point controls access to the medium by its Clients. In addition, the standard specifies a hybrid mode of operation, which combines the contention based and the contention-free medium accesses with a number of QoS specific enhancements.[9] The set of logically connected devices is known as a Basic Service Set (BSS). Multiple BSS-es are not accommodated by the 802.11 protocol, and attempts to design extensions for multiple BSS operation have not yielded practically useful results.

6.2.1 Contention-Based Medium Access: DCF

6.2.1.1 Basics

Contention-based medium access does not require a centralized control entity to control the medium access of devices in a BSS. Instead, each device within a BSS is controlled by its own distributed coordination function (DCF). DCF relies on a number of measures and mechanisms:

- slotted medium sensing
- clear channel assessment (CCA) by carrier sense and energy sense
- virtual carrier sense with the network allocation vector (NAV)
- randomized contention windows (CW) for medium access
- exponential back-off in case of transmission failure
- various types of Inter-frame Space (IFS) to prioritize medium access

Slotting in the 802.11 MAC is synchronized with carrier sense of the medium; the end of every transmission resynchronizes the slot timing of all nodes that hear it. This implies that, in the absence of medium activity, synchronization will degrade with time. Propagation delay and device implementation both play a role in the granularity of the synchronization. Propagation delay in short-range networks is not important, but in the case of outdoor wireless deployments, where the distance between two communicating devices can be more than a mile, the propagation delays can be very significant. Therefore, the slot time has to be longer; otherwise, the stations may be out of sync by one slot or more and that would reduce the efficiency of the interframe spacing used to control medium access priority for different frame types. The standard provides parameters for this purpose.

Medium sensing – called Clear Channel Assessment or CCA – is performed at two levels. True "carrier detection" is based on signal correlation detection (CD) of PHY preambles. This method is accurate and reliable. For non-802.1 signals, CCA

[9] See IEEE 802.11-2007, *Clause 9, MAC sublayer, functional description.*

is based only on energy detection (ED).[10] Because ED is much less precise and much less reliable than signal correlation based detection, the threshold for ED is 15 to 20 dB higher than that for CD. The CCA thresholds can be manipulated to optimize the system performance under various deployment scenarios; the high density deployments, for example.[11] This is not addressed in the standard.

Each 802.11 device maintains two important variables for medium access control: the contention window (CW) and the network allocation vector (NAV). A contention window is a time interval [0, CW] in units of time slots, where CW is an integer within the range of two parameters of the Contention Window CW_{min} and CW_{max}. CW determines the probability of medium collisions and therefore CW_{min} and CW_{max} should be set so as to achieve the right balance between collision probability and medium access efficiency.

In addition to the contention window, the standard provides for a medium reservation mechanism: the Network [resource] Allocation Vector (NAV). This is a timeout value that is included in each transmission frame header. The NAV interval includes the time needed for transmitting a frame and for receiving the acknowledgement. Receiving devices use the NAV value as an indicator of time period during which the device should not transmit regardless the medium status. A device transmits only when its NAV down counting reaches zero or when its carrier sense mechanism detects that the medium becomes idle – regardless of the remaining value of NAV. The combination of medium sensing, NAV timing and CW downcounting is called the Clear Channel Assessment procedure.

After a device determines the medium is busy, it performs a random back-off. It draws a random number in the range [0,CW] sets its CW-timer to this number times the slot duration. The CW timer counts down only when the medium is sensed to be idle for one or more time slots. The device will transmit only when the back-off timer expires and the medium is still idle. This is the so-called "transmit back-off." Figure 6.3 is a diagram illustrating the distributed multiple access process: Device B sends out a beacon. Its end of transmission triggers contention resolution between Devices A and C – which A wins. A's data frame contains a value for the NAV, which covers the data frame itself and the acknowledgement that is expected. After B responds with an Ack to A, C will execute the contention process which it wins. C now transmits and so the process continues.

6.2.1.2 Channel Reservation

To ensure other devices will not transmit when a device is transmitting, a device needs to reserve the channel when it is transmitting. Three mechanisms for channel reservations are available in IEEE 802.11.

---

[10] This type of sensing allows the 802.11 MAC to be spectrum etiquette compatible.

[11] See e.g. Tae-Suk [118].

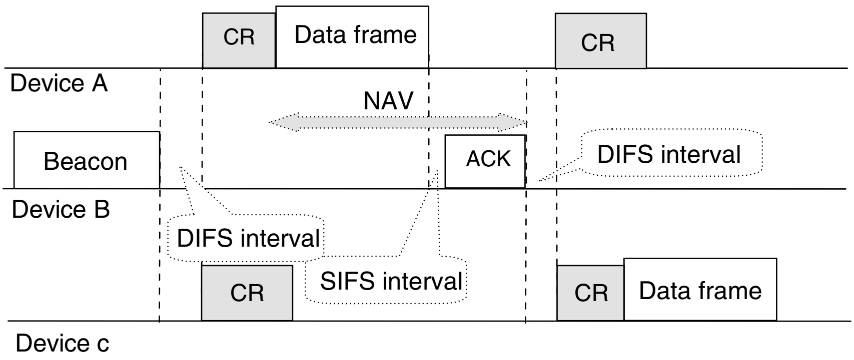

Fig. 6.3 IEEE 802.11 DCF, multiple access process

Carrier Sensing

As long as a receiver senses that the medium is busy, either with energy detection or with correlation detection, it holds off its transmissions.

NAV

Each transmission frame contains the NAV which indicates the time its transmitter will occupy the channel. Receivers defer to this.

PHY Spoofing

The transmitter reserves the channel for a longer time than required for transmitting a frame, by manipulating the values of the rate and length fields in the PHY header of the frame.

The latter two channel reservation mechanisms work by affecting the NAV values. Even in a fully connected BSS, collisions can still occur because two devices can happen to transmit at the start of the same medium access slot. The probability of this happening depends on the number of devices setting their CW value simultaneously and on the size of the CW; it is given by:

$$P_{call} = n \div CW_{min} \tag{6.3}$$

Simply extending the maximum size of the Contention Window would reduce the probability of this happening, but a large CW_{min} size, applied at every instance of medium access, is inefficient. Therefore, a balance has to be struck between the efficiency gain of shorter contention windows and the loss of throughput caused by collisions. A collision typically results in the failure of both transmissions — which wastes

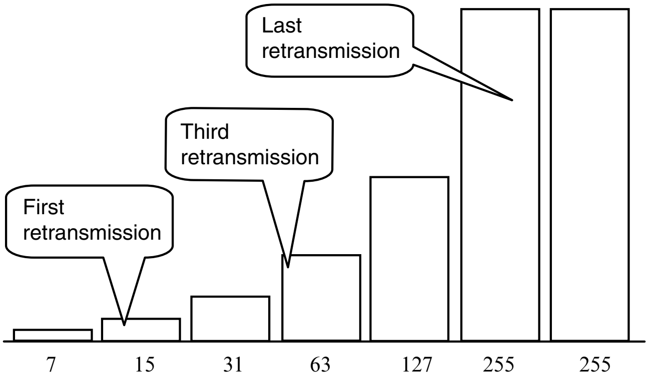

Fig. 6.4 Exponential expansion of the contention window (in slots)

channel capacity. Since the transmitters involved are not aware of the collision itself, the medium remains busy until both transmissions have terminated. Two methods are in place to recover from or prevent collisions: retry with an exponential back-off and use of RTS/CTS before a data transmission.

6.2.1.3 Exponential Back-Off

Retry of failed transmissions with an exponential back-off is a combination of error recovery and collision avoidance. The exponential back-off expands the contention window size exponentially for every transmission failure, as illustrated in Fig. 6.4. The retry and back-off can repeat until the number of retries exceeds a pre-determined retry limit.

The exponential back-off algorithm works reasonably well against collisions caused by temporary "hidden node" effects, but it is very inefficient if the frame errors are due to interference or channel impairment. However, there is no way to distinguish between these two causes of transmission failure on a per transmission basis.[12] Therefore, exponential back-off is not a good recovery mechanism in hostile RF environments.

6.2.1.4 RTS/CTS

The second method is to prevent, rather than recover from, transmission collisions caused by so-called hidden nodes: nodes that do not hear a given transmitter, but that may be heard by its intended receiver. This method uses the exchange of the control frames RTS and CTS between the transmitter and receiver, before transmitting the data frame. Figure 6.5 illustrates the use of RTS/CTS and the NAV setting.[13]

---

[12] If multiple transmission events are taken into account together, differentiation between transmission errors and collisions becomes possible.

[13] See above, Sect. 6.2.1, under *Basics.*

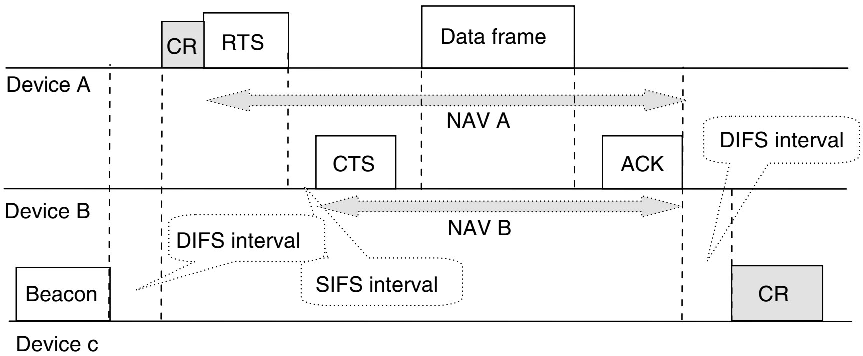

Fig. 6.5 RTS/CTS and NAV setting

The transmission cycle is initiated by A with an RTS, B answers with CTS and A sends the data frame, whereas B closes the cycle with an acknowledgement. The cycle is “bound together” by the use of the short SIFS spacing: it prevents other stations from accessing the channel – even if they did not decode the NAV value in the RTS or CTS correctly. Because these control frames are short and transmitted using a low data rate, they are more likely to be decoded successfully. If these control frames, together with the medium access overhead, are short compared to data frames, the penalty associated with the RTS/CTS procedure may be considered acceptable. Conversely, if data transmission rates are high and/or transmission frames short, the relative overhead of the RTS/CTS procedures is high and overall efficiency low. This applies clearly in case the High Throughput mode is used.

6.2.1.5 QoS

Quality of Service (QoS) is desirable for latency sensitive applications such as Voice over IP and Video over Wireless. The IEEE 802.11 standard specifies an enhanced distributed channel access method (EDCA) to accommodate the needs of QoS sensitive applications. EDCA provides four access categories (ACs) for traffic with various latency requirements: background, best effort, voice and video. Each AC maintains its own DCF instance with separate EDCA parameters and transmission queues. The EDCA parameters include the Arbitration Inter-Frame Space (AIFS), CW_{min}, CW_{max}, TXOP[14] window size and retry limits. Different levels of service are created by differentiation in these parameters. Each access category may be given a different A-IFS; the relative lengths of the A-IFS values determine the relative priority of medium access among frames *queued within a station* for transmission. EDCA does not

[14]Under the rules of the EDCA and HCCA, a client can request the Access Point to reserve the medium for up to 8,192 μs.

provide a means to order the transmissions of stations. EDCA can be used in many ways. For example, a voice frame will deterministically win the contention for the medium over a best effort frame when the AIFS for the Best Effort access category is set to be greater than the sum of CW_{max} and AIFS for the voice access category.

At the level of the BSS, this method of traffic flow differentiation is effective *only* if the capacity of the channel exceeds the offered QoS traffic load of all devices. Further channel access under EDCA is statistical rather than deterministic. This has the advantage of avoiding the unpredictable behavior associated with contention free medium access in a shared medium.

6.2.2 Contention-Free Medium Access: PCF and HCF

Contention-free medium access relies on a centralized entity. In the IEEE 802.11 standard, the point coordination function (PCF) provides coordination of the frame transmissions by multiple devices. Implementation of a PCF is optional; few products on the market actually implement it. Typically, the PCF resides on the Access Point of a BSS. The PCF relies on a number of essential features:

- A method to schedule the Access Point's transmissions and client transmissions. It decides which device has the right for channel access at a given time, and how long it can occupy the channel. The standard does not specify an algorithm or process for this.
- A mechanism to collect client status information and to distribute the scheduling information.
- An error recovery mechanism in case the scheduling information fails to be delivered timely due to causes such as interference, overly occupied channel or out-of-date scheduling information arising from changes in the transmission rate in response to channel changes. The standard does not define this mechanism.
- A mechanism to handle the overlapping BSS'es[15] where multiple scheduling entities co-exist. The standard does not define this mechanism.
- A mechanism to handle co-existence with non-compatible devices. The standard does not define this mechanism.

During PCF operation, the Access Point polls its clients one by one with CF-Poll frames. A client transmits frames to the Access Point only when it receives a polling frame. Using a special frame header, the Access Point informs clients in advance of the time of transmission for each of them.

The medium access rules during PCF operation are a subset of the DCF rules: short interframe spaces (SIFS) are used to assure that only devices in PCF mode have access to the medium. The transition from DCF to PCF operation is enabled by

[15] Much effort has been invested in developing such a mechanism, but success has been elusive. This should come as no surprise: the coordination effort has to take into account an infinite series of potential interfering BSSes. Operating each BSS on its own channel avoids this problem.

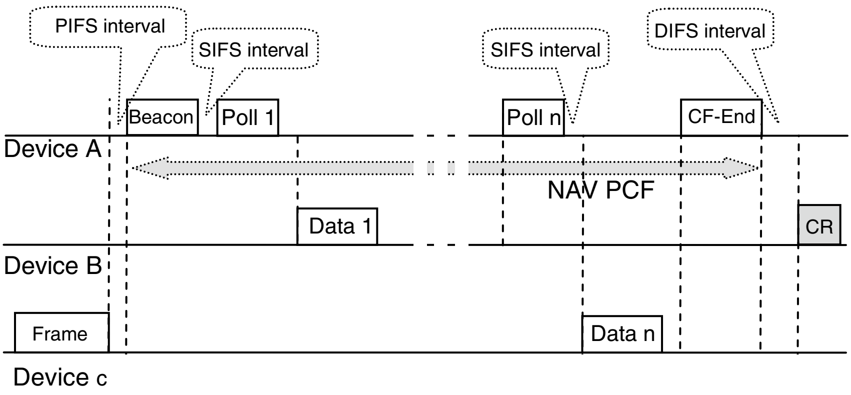

Fig. 6.6 Example of PCF operation

using a shorter inter-frame space, the PIFS. This is how the PCF ensures the co-existence with non-PCF devices. Figure 6.6 shows an example of PCF frame transfer initiated by Device A, which pre-empts all other devices by means of the PCF Interframe Space. The duration of the PCF period is signalled in the NAV of the beacon sent by A at the start of the PCF period. During this period, all exchanges are spaced with the SIFS interval, so as to prevent non-participating devices to initiate the contention resolution process.

The Hybrid Coordination Function specified in the IEEE 802.11 standard is a hybrid of PCF and DCF. In HCF Coordinated Channel Access (HCCA), the contention-free period (CFP) and contention period (CP) do not alternate; instead, the latter can be started at any time. DCF rules apply during the contention periods and PCF rules apply during the contention-free period. The contention-free period is reserved by the NAV mechanism in the beacon frames transmitted at the beginning the contention-free period. CF-END is used by the PCF to end a contention-free period and reset the NAV values for all devices.

In theory, contention-free medium access allows for a higher protocol efficiency and a higher network capacity than contention-based medium access. It does not have the overhead of back-off and avoids the need for CW size increase with increasing number of stations. In practice, however, these benefits can hardly be fully achieved, notably on a busy channel. First of all, the PIFS interframe space is sent out only by one device – typically, the access point. Devices that do not recognize the start of the PCF will disrupt it. Further, interference can easily disrupt the PCF scheduling and retransmissions tend to negate the effectiveness of the PCF for QoS sensitive applications – its major justification.

Given the complex interactions of medium, protocol, traffic flows, interference, the task of scheduling PCF operations is very challenging,[16] even without overlapping BSS

[16]In licensed systems, such central coordination is less complex because there is very little unpredictable interference.

considerations. This is one of the major drawbacks for PCF operation. Because few – if any – devices implement the optional PCF, PCF implementations have to coexist with DCF in a BSS, which makes the PCF even less flexible and less efficient. A number of references[17] indicate that PCF performs poorly either alone or mixed with DCF.

6.3 Adaptive Medium Access Control

6.3.1 Introduction

The principle of listen-before-talk in medium access control has proven to be both versatile and useful, but not necessarily very efficient or robust in the face of adverse propagation conditions or interference. This is also true for the medium access methods defined in IEEE 802.11. These access methods implicitly assume near perfect RF connectivity, which assures that *"all-hear-all."* Transmission failures are assumed to be either an indication of insufficient SNR at the receiver or the result of hidden node interference. That latter is counteracted by the exponential back-off mechanism. A lower data rate reduces the SNR required and, therefore, frequent transmission failures cause a transition to a lower data rate. However, when the transmission failure is due to burst-like interference at the receiver, switching to a lower data rate does not help. A lower data rate tends to congest the medium even more, causing more collisions and even lower data rates, possibly ending in starvation.

6.3.2 Input Requirements

Adaptive medium access requires statistical data on the channel conditions, the methods to measure them and the means to communicate the related feedback data from the receivers to the transmitters. Many network management related parameters[18] defined in the IEEE 802.11 standard can be used for this purpose. For example, a minimum set of statistics could include:

6.3.2.1 Frame Success Rate at the Transmitter

The success rate for a given receiver, as seen at the transmitter, reflects the local channel conditions *and some aspects of* RF channel conditions. It can be measured by observing, for each receiver, the ratio between frames transmitted and acknowledgements received. The success rates can be further categorized by frame size and data rate. Such a measurement is blind to the causes of receive failure: a low SNR may result in

[17] Lindgren et al. [85] and Barry et al. [19].

[18] These have been developed in 802.11 Task Groups 11 k and 11v.

the same receive failure rate as a low SIR. Collisions at the receiver in an overloaded wireless LAN can similarly degrade the throughput even with an adequate SNIR, but the impact should vary with frame duration (= frame size × data rate).

6.3.2.2 Frame Success Rate at the Receiver

The success rate seen at the receiver for a given transmitter and receiver pair represents the local channel condition at a receiver; however, it is dominated by the background interference at the receiver. The frame success rate can be categorized by frame size and data rate. The observed values reflect both the actual SIR and the multi-path conditions of the RF channel. In general, the instantaneous receiver success rate is not a good predictor of future receive success. Therefore, receiver success must be observed over a certain period of time to smooth the short-term variability. Averaged over many transmitters, this abstracted receiver success rate reflects local interference levels and, therefore, it is an adequate predictor of actual receive success. Different success rates for different packet durations are helpful to differentiate between static channel parameters and dynamic interference.

6.3.2.3 Channel Occupancy

Presence of strong signals from nearby transmitters could be considered foreground interference or, more naturally, channel occupancy. It may be very different from station to station. The CCA[19] statistics of the RF transceiver over some period of time reflect the local channel occupancy. Channel occupancy at the transmitter indicates the probability that a transmission will cause a collision in the vicinity of the transmitter. Channel occupancy at a station indicates the probability that receiving a frame will fail due to local interference. Therefore, it is advisable to use the maximum of the channel occupancy values observed at receiver and transmitter, in determining such parameters as contention window size and packet duration.

6.3.2.4 Background Noise Level at the Receiver

Whereas background noise plays no role at the transmitter, it may be the major determinant of successful receive operations for a given transmitter. Transmissions from a nearby transmitter may not be affected, those of a further away transmitter may have to be scaled back in transmission rate. However, transmission failure due to a high background noise receiver at the receiver need not lead to exponential back-off and, therefore, having this information is valuable to a transmitter. Given that the background noise level will not fluctuate much, a station could broadcast this information at a slow rate.

[19] See Sect. 6.2.1, under *Basics*.

Communicating feedback from the receiver to the transmitter requires only a few bytes of data. The frequency of feedback exchanges should be high enough to reflect the actual behavior of the channel, but it need not be continuous: the short-term variability of the channel is too high to warrant real-time feedback. The feedback data can be sent as a separate management frame, but including it in the header of a data frame is much more efficient. A separate management frame not only takes far more channel time, it also adds to the channel load and interference probabilities for other stations.[20]

6.3.3 Adaptation Behavior

Adaptive medium access has been addressed by a number of research projects[21]; these have identified various means of adjusting medium access parameters.

Adaptive medium access can make use of the above information to adjust its medium access parameters, e.g.:

- Contention Window size
- Transmission retry limit
- Transmission data rate
- Frame Fragmentation or Aggregation thresholds
- Back-off algorithm (flat or exponential)

The ways in which these parameters may be modified in response to channel information are many. Here only a few examples are given.

6.3.3.1 Contention Window Size

CW size should increase with the observed packet error rate, as well as with the observed channel occupancy at a given receiver. By varying the CW size and keeping track of the observed receiver success rate, the optimum the probability of transmission attempts for a given receiver can be determined.

6.3.3.2 Transmission Retry Limit

The transmission retry limit could be increased to counter burst-like interference at the receiver and reduced if channel occupancy at the receiver is high.

6.3.3.3 Transmission Data Rate

The transmit data rate adjustment makes sense in response to changes strong (hidden node) interference.

[20] See Camp [31].

[21] See e.g. Zhu [129].

6.3.3.4 Back-Off Algorithm (Flat or Exponential)

If the statistics indicate that the dominant cause of transmission failures is excessive interference, the CW size for transmission retries should be kept constant rather than increasing, as in conventional exponential back-off.

6.3.3.5 Frame Fragmentation or Aggregation Thresholds

Frame fragmentation can help in countering the effects of repetitive interference, e.g. as caused by microwave ovens. Frame aggregation is beneficial at high data rates, but only if the background interference level at the intended receiver is low.

In summary, adaptive medium access effectively combats the performance degradation in difficult RF conditions. It provides a means to maximize the performance of the IEEE 802.11 MAC, and by extension the performance of other types of medium access protocols based on the listen-before-talk principle. The success of an adaptive medium access design relies on many implementation considerations such as the selection of the statistics, the methods to communicate receiver feedback and the algorithms to apply the statistics adaptively. All of these leave a lot of room of innovation for the device, without compromising compatibility with the IEEE standard.

6.4 Medium Access Control for Directional Antennas

6.4.1 Basics

Directional antenna technology has been developed for more than 70 years. Directional antennas provide an extended operating range and better spatial reuse.[22] They have been adopted in many wireless communications systems. Examples are the third-generation cellular systems (e.g. 3GPP Release 6) and the Adaptive Antenna System (AAS) of IEEE 802.16. The same benefits can be realized in other systems such as IEEE 802.11 based mesh networks.

A directional antenna may be implemented as an antenna array, or multi-element antenna combined with the *beam-forming* technology – which may be analog or digital in nature. The latter can be seen as a form of multiple input-multiple output (MIMO) technology. An antenna gain pattern is generated by applying various delays and weighted summation over the outputs of the antenna elements.

Depending on the flexibility of the antenna gain patterns, a directional antenna can be classified as a sectored antenna or dynamic beam-forming antenna. A sectored or switched antenna produces a fixed, pre-defined antenna gain pattern. A beam-forming antenna, on the other hand, may be used to generate a gain pattern specific for the RF propagation path between the transmitter and the receiver.

[22] See Chap. 4, Sect. 4.2.2 for more details.

Dynamic beam-forming antennas are capable of beam-forming gains, null steering, spatial multiple reuse and spatial diversity with the combination of space-time codes. They can be used at transmitter and receiver and, as a result, such antennas possess a number of desirable benefits, for example:

- Higher SNIR through beam-forming gain
- Lower interference through null steering
- Higher spectral reuse with spatial multiple reuse
- More robustness to multi-path fading through spatial diversity combined with space-time codes.

The adaptive medium access as described above operates in the time dimension and it assumes that the antennas are omni-directional. Medium Access Control with directional antennas operates in both the time and spatial dimension. In the following sub-sections, we will use the application of directional antennas in Wi-Fi mesh networks, as an example to discuss the benefits and challenges, especially with regard to medium access control, of using Directional Antennas in a shared RF environment.

6.4.2 *The Application of Directional Antennas in Mesh Networks*

Mesh networks are constrained by issues in spatial reuse, routing performance and power efficiency. Because of these issues, directional antennas have been widely explored for their potential to improve outdoor mesh networks. Practice as well as research shows that Wi-Fi mesh networks, more so than the conventional single hop infrastructure Wi-Fi networks, are likely to exhibit undesirable behavior such as "node starvation," "unfairness,"[23] "TCP instability,"[24] excessive hidden node interference and exposed node effects. Using a directional antennas to address these issues raises interesting challenges.

6.4.2.1 Benefits of Directional Antennas

Directional antennas greatly improve signal quality when the transmitter, the receiver, or both, are so equipped. Dynamic transmit and receive beamforming[25] can also be used to improve the SIR of a link, by forming nulls in the direction of interference sources. These improvements in signal quality translate into higher data rates and/or better operating range. In mesh networks, a longer operating range means a lower number of hops between two network nodes and, therefore, lower frame delays.

[23] Fairness in the context of medium access is a very difficult subject – if only because of the many ways in which medium utilization can be measured. See chapter 12.2.5, *Fairness.*

[24] See Xu and Saadawi [126].

[25] With such antennas, SNR improves linearly with the number of antenna elements. See also Liu and Li [87].

Improved spatial reuse is another desirable benefit provided by directional antennas. They also provide a solution to the “exposed node issue[26]” by allowing for multiple flows to be active and independent at the same time. Such improvement in spatial reuse may multiply the achievable throughput. This has been confirmed in research and experimental work,[27] which showed that directional antenna technology provides significant improvement in throughput and spatial reuse.

The use of directional antenna in wireless mesh networks is not without difficulties, notably with regard to medium access. In the following, commonly encountered challenges in the PHY layer, the MAC sub-layer and Routing layers are identified. Some of them remain unsolved.

6.4.2.2 The PHY Layer and Directional Antennas

Operations of LBT-based systems like Wi-Fi rely heavily on physical carrier sensing. All listen-before-talk medium access mechanisms assume omni-directional antennas, which are necessary to assure that “*all-hear-all.*” However, beam steering and antenna switching reduce the directions as well as the time window during which the receiver will receive signals from other transmitters. The implication is that the standard approach to Clear Channel Assessment of IEEE 802.11, which is based on decoding the transmission header, will not work. Medium activity may not be relevant to the assessment of a transmit opportunity: the detected signal may come from the intended receiver of a transmission; in that case, waiting for a clear channel indication makes sense. However, the detected signal may also come from other sources. In case the intended receiver also uses a directional antenna, it may not even see these other sources and, therefore, the transmitter could proceed with transmission. Differentiation between these two cases requires special mechanisms in the MAC sub-layer.

6.4.2.3 MAC Protocols and Directional Antennas

Common issues in wireless networks, such as hidden nodes issue and exposed nodes, are exacerbated when directional antennas are used. The conventional medium mechanisms such as RTS/CTS and back-off with expanding contention window sizes have no effect or worsen things. Therefore, a “directional medium access” approach is needed that takes care of these issues and that is effective in a homogenous network as well as a heterogeneous network. In the former case, all network nodes use the same directional antennas; in the latter, both directional and non-directional antennas are used. The following assumes that each node uses the same antenna direction for transmission, receiving and medium sensing.

[26] An “exposed node” is a node prevented for transmitting, although that transmission would not interfere with the transmission for which it has to defer.

[27] See Ramanathan [105] and Spyropoulos and Raghavendra [114].

As noted above, medium sensing with directional antennas can only provide information about activity on the channel in a given direction. In a homogeneous network, the intended receiver also uses a directional antenna and it may not see other transmitters. Therefore, a transmitter could proceed with transmission even if CCA indicates the channel is busy. Distinguishing between the activity of the intended station and other stations requires additional information such as the known direction of other antennas in the vicinity of the remote station or estimated times of its transmissions. The latter calls for some form of synchronization between the nodes of such a network, so that nodes can schedule their transmissions and/or antenna pointing correctly.

6.4.2.4 The Deafness Issue

In a heterogeneous network, a well-known and significant issue resulting from the use of directional antennas is the so called "Deafness issue." The Deafness issue can be present in many forms. In the example illustrated in Fig. 6.7., Node A, Node B, Node C and Node D are not hidden from each other if they all use omni-directional antennas. When A communicates with B using a directional antenna and other nodes use omni-directional antennas, A becomes deaf to the transmission of C due to the direction of its antenna gain. Because of the direction of transmission, C cannot hear the transmission of A to B either. C may transmit to D while there is on-going transmission from A to B. As a result, there may be a collision at B, resulting in failure of a transmission from A. Note that even if C or A transmits RTS before transmitting a frame, such collisions cannot be avoided. If C transmits RTS to A while A is transmitting to B, C will not receive a CTS response from A. It will therefore keep on retrying with larger and larger contention window sizes before the transmission finally times out. In that case, not only C consumes much channel bandwidth without success, D will also be adversely impacted by being deferred unnecessarily from medium access due to channel reservation by C.

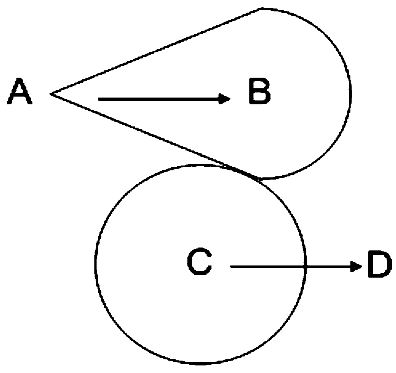

Fig. 6.7 Deafness issue

In many network topologies, especially in the case of wireless mesh networks, the deafness issue can be created by heavy shadowing of transmitters as well as by the use of directional antennas on some but not all stations. In either case, deafness causes other problems such as unfairness or even deadlock which cripple a network. These issues are not easily solved with conventional MAC mechanisms.

Deafness also occurs in a homogeneous network if directional antennas are switched in time as well as space. In such an environment, busy tones may not work and more explicit means of coordination are required. Centralized antenna scheduling does not scale well and it is not robust: errors in directions and timing can easily wreck havoc. On the other hand, a distributed antenna scheduling scheme has major issues too: optimal scheduling requires knowing the traffic levels of adjacent network nodes, and the same applies for the adjacent nodes themselves. Dropping the notion of "optimal" scheduling allows an approach in which nodes exchange traffic flow information and adjust their antenna schedules accordingly. Such an approach would work only for homogenous networks.

In addition to the approaches focusing on *avoiding* deafness issues in heterogeneous networks, some have looked into the possibility to *detect* deafness and adaptation to reduce its impact.[28,29]

Enhancement of conventional medium access mechanisms to address the issues encountered in using directional antennas has long been an active research topic. Solutions based on exchanging information between the nodes in a distributed manner[30] have been proposed. Information, in some cases as simple as a busy tone, is exchanged omni-directionally before or after a directional transmission. Although in theory those solutions could alleviate the adverse effects of directional antennas, successful implementations have many difficulties to overcome. For example, the busy tones may be hard to identify as corresponding to a transmitter and receiver pair.

6.4.2.5 Routing and Directional Antennas

Directional antennas add more complexity to the already difficult problem to build optimal mesh routes. For example, broadcast messages are often used in discovering routes. The delays caused by directional transmissions of these broadcasts may result in sub-optimal routes, since the discovery requesting messages traversing sub-optimal paths can trigger route reply messages long before an optimal requesting message arrives. Time stamps involved in routine protocols might help to mitigate the issue. Moreover, while the destination transmits a directional reply, it may not be able to receive more optimal request messages. These side effects of directional

---

[28] See Neufeld and Grunwald [99].

[29] The deafness issue may be considered a special case of RF coupling asymmetries that can interfere with protocol effectiveness. See also Chap. 7, Sect. 7.2 *Legacy Wireless LANs*.

[30] See Choudhury and Vaidya [39] and Huang et al. [69].

antenna usage reduce the value of conventional routing approaches, notably the dynamic one like AODV which rely on network wide discovery broadcasts. Other routing protocols that leverage the presence of gateway nodes to focus path discovery will perform better over directional antenna links.

In summary, with suitable adaptations in medium access control, directional antenna technology clearly has the potential to increase network capacity multiple times through increased spatial reuse and to improve end-to-end routing performance via fewer hops in mesh networks. However, this potential will not be fully realized before a number of challenges – including relaxation of regulatory limits on radiated power – are overcome.

6.5 OFDMA: Multiple Access with OFDM

OFDMA stands for Multi Access OFDM, a technology that has been adopted in a number of recent wireless standards such as ETSI DVB-RCT, IEEE802.16a, IEEE802.16d and IEEE 802.16e. OFDMA combines shared use of the frequency and time dimension to optimally fill the available capacity in both the down link and the uplink. OFDMA is being considered for Wi-Fi technology (under heading of IEEE 802.11 ac).

Conventional TDMA provides time to multiple users in some order, possibly sequential. OFDMA goes beyond that by also selectively allocating sub-carriers within OFDM symbols to users in the frequency domain at the same time. In OFDMA, an RF channel is treated as a collection of subcarriers grouped into multiple sub-channels, each composed of multiple, non-adjacent carriers. One or more sub-channels are designated to each user. See Fig. 6.8. Adaptive Modulation and Coding per Sub-Channel can be applied to further enhance performance.

The combination of TDMA and FDMA provides very high efficiency in channel utilization. In addition, OFDMA allows for burst structures of both uplink and downlink to selectively increase system throughput for some network nodes. The burst structure for a given user or client device is defined by one Sub-channel in the Frequency domain and a number of OFDMA symbols in the time domain. Table 6.1 shows an example of uplink and downlink OFDMA specification.

OFDMA possesses a number of distinguishing advantages. It effectively handles interference from both the neighboring cells and from within the same cell with carrier permutations. OFMDA improves diversity in three dimensions – space, time and frequency – for a better connectivity. Spatial diversity is achieved by using antenna diversity for OFDMA devices. Frequency diversity in OFDMA is provided by spreading carriers over the used spectrum. Time diversity is enabled by interleaving of carrier groups in time. OFDMA may be combined with forward error correction to improve transmission over noisy or weak sub-channels. In addition, uplinks for multiple users are made orthogonal with synchronizations in both time and frequency. OFDMA also mitigates multipath effects without using equalizers and training sequences.

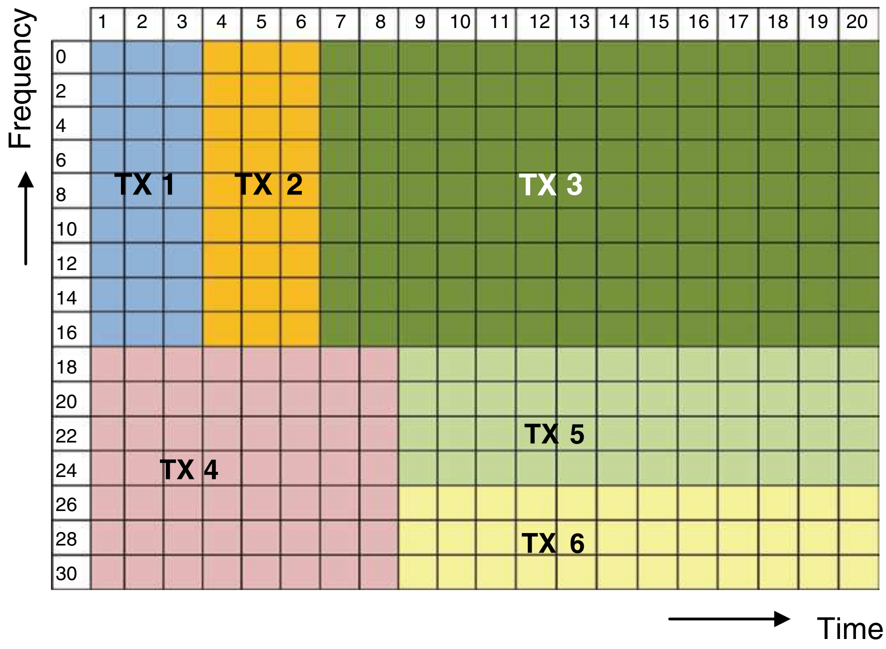

Fig. 6.8 OFDMA with TDMA and FDMA components

Table 6.1 Example of OFDMA specifications

	Uplink	Downlink
FFT size	2048	2048
Guard intervals (symbols)	1/4, 1/8, 1/16, 1/32	1/4, 1/8, 1/16, 1/32
Coding mandatory	Concatenated RS GF(256) and convolutional coding	Concatenated RS GF(256) and convolutional coding
Coding optional	Convolutional Turbo Code (CTC), Turbo Product Code (TPC)	Convolutional Turbo Code (CTC), Turbo Product Code (TPC)
Modulations	QPSK, 16QAM, 64QAM	QPSK, 16QAM, 64QAM
Sub-channels	32	32
Carriers per sub-channel	53	53
Pilots per sub-channel	5	5
Data carrier per sub-channel	48	48
Symbols per burst	3	Multiple
Sub-channels per user	1–32	1–32

OFDMA systems typically provide adaptive modulation e.g. QPSK, 16QAM, 64QAM and 256QAM. Adaptively using the highest modulation significantly increases the cell capacity. It accommodates QoS as well, e.g. by using small bursts of around 100 symbols per user for a better statistical multiplexing and

reduced jitter. All these desirable properties make OFDMA a very useful technology for advanced wireless communication systems that have to share scarce frequency bands efficiently.

OFDMA has drawbacks including increased sensitivity to frequency offset, phase and reduced resistance to frequency selective fading and complex signal processing. However, advances in semiconductor technology have reduced the impact of these drawbacks. Therefore, OFDMA is used in the later generations of WiMAX systems, as well as in LTE – the long-term evolution of third generation cellular systems.

6.6 Bluetooth and IEEE 802.15

6.6.1 Overview

Bluetooth is an example of another successful wireless commodity technology that has become widely used in a variety of applications. Bluetooth is the trade name for a personal area network technology that supports a “reliable” service for voice communications, as well as a variety of formats for data transmission. It uses the 2.4 GHz ISM license-exempt band for interconnecting computing and communication devices and accessories. In the United States and most other countries, this band is divided into 79 channels, each with a width of 1 MHz – the value used as the bandwidth for the frequency hopping transceiver of Bluetooth. By creating separate hopping patterns, multiple “piconets” can be created which share spectrum without interfering with each other. This is the basic concept of the Bluetooth design, which eventually became the IEEE 802.15.1 standard for wireless personal area networks (wireless PAN).

The Bluetooth protocol stack does not conform to either the OSI network stack model or IEEE 802 stack model. IEEE 802.15 managed to adapt the Bluetooth’s protocol stack into the IEEE 802 model. The three lowest layers of Bluetooth protocol stack roughly correspond to the physical and MAC layers of the IEEE model.

Since IEEE 802.15.1, the first wireless PAN standard, was approved by the IEEE in 2002, a number of Bluetooth specification revisions have been developed. Bluetooth devices provide various gross data rates, depending on their standard versions, as shown in Table 6.2.

Three power classes are specified for Bluetooth devices. Their transmission powers and ranges vary with their power classes. A Bluetooth device usually supports only one of three classes, typically the low-power and short-range one, as shown in Table 6.3.

The implementation of a complete Bluetooth design is intended to be possible on a single chip, which assures low cost, usually well under $5. Thanks to this low cost, Bluetooth has been widely adopted for many low-cost, short-range, and low-power applications, such as mobile phone headsets, cordless telephony, PDA-computer synchronization.

Table 6.2 Bluetooth versions and supported data rates

Version	Gross data rate (Mbps)
1.2	1
2.0+EDR	3
3.0+HS	24

Table 6.3 Bluetooth power classes

Power class	Power (mw)	Power (dBm)	Range (meter)
3	1	0	10
2	2.5	4	20
1	100	20	100

The Bluetooth SIG specifications and the 802.15 WPAN versions are not identical. The Bluetooth SIG specifications cover a complete system from PHY to application layer, while IEEE 802.15 is limited to the physical layer and the data link layer. Bluetooth specified 13 specific applications, called profiles, to be supported with different protocol stacks for each. This introduced tremendous complexity and resulted in the first Bluetooth specification, the V1.0, with 1,500 pages. More details are available from the Bluetooth SIG.[31]

6.6.2 Medium Access Mechanisms

The basic Bluetooth network is called a piconet, which contains one master node and up to seven active slave nodes. Up to 255 non-active nodes, the so-called parked nodes, are allowed in a piconet. Multiple piconets can be co-located or overlap. When multiple piconets are connected with bridge nodes, they form a scatternet.

A piconet is a centralized TDD (Time Division Duplex) system, using a scheduled transmission scheme with time slots. The master node controls the clock and dictates which slave node has the right to communicate at a given time slot. The master and slave nodes take turns to transmit. Slave nodes are typically dumb and constrained by low-cost design. They just do whatever the master tells them to do. Slave nodes are not allowed to directly communicate with each other.

The core of the Bluetooth medium access design is its frequency hopping radio with a hop rate of 1,600 hops per second. Bluetooth employs frequency shift keying with 1 bit per Hz, which results in a gross data rate of 1 Mbps. However, the payload data rate is usually much lower because of overhead: settling time of 250–260 μs for the transceiver after each hop and the overhead for the access code and frame header of 126 bits.

[31] See http://www.bluetooth.com/English/SIG/Pages/default.aspx

The hop rate yields a "dwell time" per frequency channel of 625 μs, which is the basis of time slots for a Bluetooth piconet. All nodes in a piconet simultaneously hop over a sequence of frequencies dictated by the master. The fast hopping and use of narrow band transmission helps to reduce mutual interference. The sequence in which frequency slots are used form a so-called hopping pattern. Each piconet has its own hopping pattern; provided the number of hopping patterns is limited, the probability of collision – i.e. the simultaneous use of the same frequency slot by two piconets – is low to negligible. By having nodes participate in multiple piconets, a so-called "scatternet" is created. A slave can be a master in another piconet and this allows interference among piconets to be avoided.

The hopping patterns are overlaid with a small set of dedicated services for voice and data. Hop sequences are determined by a master node that tells slaves what hop frequencies to use.

6.6.3 Bluetooth Traffic Patterns

Typically, Bluetooth is used as a personal area network – people use it to connect gear they use in close proximity to themselves. Bluetooth systems have two dominant traffic types: the Synchronous Connection Oriented type (SCO) and the Asynchronous Connection oriented Link type (ACL). Both are an attempt to create predictable communication services. The former is used to transport voice and other interactive services; the second is used for data. Whereas the SCO traffic is bi-directional on a per slot basis, the ACL traffic pattern allows multiple slots to be concatenated so as to create a more efficient transport mechanism. On the other hand, whereas the SCO pattern is constant for a duration of a connection (e.g. a call), the ACL pattern varies with the offered load: a file transfer will look like asymmetrical SCO pattern with all slots occupied, whereas a mouse or keyboard link will show only intermittent activity (Figs. 6.9 and 6.10).

6.6.4 Interaction with Other Spectrum Users

The issue of sharing spectrum with other systems was not a major concern at the time that the Bluetooth specification was developed. This was partially due to the fact that there were few other users of the 2.4 GHz band at the time. In addition, the fast hopping pattern and the low power appeared adequate minimize interference between the piconets of different users.

Later, the issue of sharing with fast growing "802.11" was recognized, but it took until the early years of this century for a formal publication on the matter.[32] It did

[32] See 802.15.2-2003-Part 15.2: Coexistence of Wireless Personal Area Networks with Other Wireless Devices Operating in Unlicensed Frequency Bands.

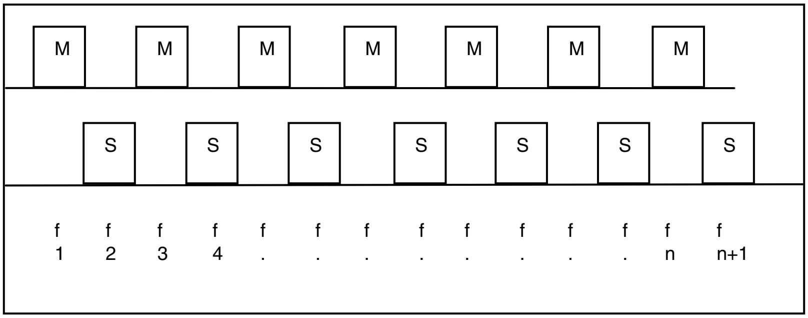

Fig. 6.9 Bluetooth traffic example: SCO service

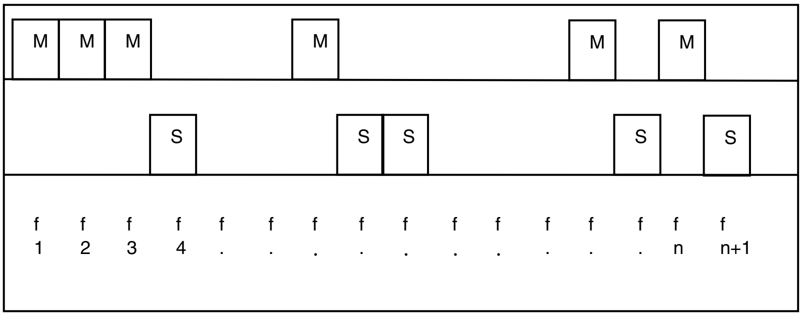

Fig. 6.10 Bluetooth traffic example: ACL service

not solve the problem, but it pointed towards a possible solution: some form of tight coupling that would allow both systems a fair degree of autonomy while reducing the mutual interference. The even more effective solution of "moving" 802.11 based systems to the 5 GHz band never got much traction.

It will be clear from the preceding that the rules for frequency hopping (at least 75 hops for 1 MHz wide hoppers) actually forces interference to occur. Typically a wireless LAN occupies 20 MHz of spectrum (nominally 16.5 MHz for OFDM and 22 MHz for CCK). Allowing a Bluetooth picocell to adjust its hop pattern to avoid colliding with the wireless LAN transmission would benefit both systems.

In 2002, the FCC authorized the use of 15 hops for wideband hoppers (those ranging from 1–5 MHz) in the Second Report & Order of May 16, 2002 (FCC 02–151). The key statement in this Report & Order is:

> *Additional actions taken today will permit the use of as few as fifteen hopping channels for FHSS in the 2.4 GHz band. These systems will be able to use channel bandwidth up to 5 MHz wide, but they must reduce their output power to 125 mW if fewer than 75 hopping channels are used. This action will allow new FHSS systems to better avoid interference*

than today's systems by enabling them to avoid occupied channels. The Commission also eliminated the processing gain requirement for DSSS systems, concluding that manufacturers have market-driven incentives to design products that they can withstand interference from other radio frequency devices.

While doing so, the Commission assured the opportunity for adaptive hopping by permitting those systems to use as few as 15 hopping frequencies for the express purpose of avoiding interference. This modified rule allows Bluetooth systems to avoid colliding with a nearby wireless LAN. This is implemented through adding a table with available and non-available RF channels and modifying the basic channel hopping sequence, so that unavailable RF channels are remapped uniformly onto available RF channels; this maintains the pseudo-random nature of the hop sequence, while allowing both master and slave to generate the correct hop frequencies. However, this modified hop sequence carries a price for the interference reduction: the frequency space available for orthogonal (= independent and therefore non-colliding) hop sets is reduced and, therefore, the number of independently operating piconets is reduced as well. In many cases – e.g. a few Bluetooth devices on a desktop – this will not be a problem.

More detail on this subject is provided in Chap. 8, Sect. 8.1.

6.7 IEEE 802.16 and WiMAX – Centralized Medium Access

IEEE 802.16, the "Air Interface for Fixed Broadband Wireless Access System," is the IEEE standard for Wireless Metropolitan Area Networks (WMANs). WiMAX, Wireless Microwave Access technology for broadband Wireless Access, is the name of an industrial consortium formed to promote wireless technology and certify the devices for interoperability.

WiMAX technology is based on the IEEE 802.16 series standards, is emerging as a major player[33] in commercial wireless networking. Its applications include fixed point to multipoint networks – i.e. wireless DSL as well as mobile, i.e. cellular communications. In addition to its competition with Wi-Fi in the outdoor environment, WiMAX was considered by some as a threat to 3G cellular technology, because its proffered data transfer rates would meet targets of LTE (Long Term Evolution), the new generation cellular standard.

The 802.16/WiMAX air interface has evolved considerably over the years and it is worthwhile to consider it in view of potential adjacent band interference issues with license exempt Wi-Fi networks. This is taken up in more detail in Chap. 9, Sect. 9.2.

[33] Some saw it as yet another wireless standard that was destined to join other little used standards in the background of the industry standards scene.

6.7.1 Overview

The IEEE 802.16 standard was originally developed to solve the “last mile” issue, i.e. as a wireless replacement for cable and (A)DSL it was limited to fixed operations. Development of the standard started in July 1999 and the first 802.16 standard was approved in April 2002. A number of revisions, 802.16a, 802.16b and 802.16c were released to address certain spectrum issues and other enhancements. 802.16d, also known as 802.16-2004, is a major revision of 802.16; it focused on fixed applications and was released in 2004. Since then, the scope of 802.16 has been extended to nomadic and mobile use. In 2005, a new flavor of 802.16, the 802.16e, also known as 802.16-2005, was released, it enabled full nomadic and mobile use.

Table 6.4 summarizes the major characteristics of 802.16d and 802.16e.

A number of amendments have been developed to cover other aspects such as management (802.16 g), coexistence mechanisms for license exempt operation (802.16 h) and, as of 2010, very high speed operation and simultaneous multiple client access using OFDMA (802.11 m). The latter is candidate for the next (4th) generation of cellular technology under the heading on IMT-Advanced. The early versions of the IEEE 802.16 standard were not specific with regard to frequency of operation. Later, when cellular applications became a priority, the focus shifted to licensed frequency bands below 6 GHz, which now include the 700 MHz UHF band and TV White Space spectrum.

Nonetheless, there has always been interest in deploying this technology in license exempt spectrum. Notably, IEEE802.16d has been designed with unused spectrum of 10–66 GHz range in mind. Millimeter waves have the property to be strongly absorbed by water, therefore by rain, snow and heavy fog. As a consequence, the IEEE 802.16 standard provides for more error handling than the IEEE 802.11 standard, which primarily targets indoor wireless networks. 802.16 employs forward error correction with Hamming codes, whereas 802.11 uses only a checksum – for error detection. Further, the standard provides for multiple directional beams, which are useful to improve the channel utilization. These and other features of the standard facilitate deployment in license exempt spectrum, such as the 3,650 MHz BWA band in the US and the 5.8 GHz BWA band in Europe. In both cases, these systems may be subject to interference from other users – which include high powered Wi-Fi systems. However, the applicable

Table 6.4 Major characteristics of IEEE802.16d and IEEE80216e[a]

Standards	Uses	Rates (Mbps)	Range (km)
802.16d	DSL/cable replacement	Up to 75	Up to 75
802.16e	nomadic/mobile	Up to 15	2–4

[a]The table must be read with some care: the highest data rates do not go with the largest operating distances.

regulations assume that the operators will resolve such interference issues on a bi-lateral basis. For that purpose, registration data bases have been established in the US and in Ireland.

6.7.2 TDD versus FDD

Two modes of operation are supported: FDD (Frequency Division Duplex) and TDD (Time Division Duplex). The former goes well with streaming services like voice and video, and the latter promises to handle data heavy traffic flows more efficiently.[34]

For fixed applications, TDD/TDMA is a perfect solution because the directional antennas of the subscribers provide significant isolation from neighboring cells. This isolation reduces the re-use distance between two cells using the same frequencies. However, nomadic and cellular systems do not use directional antennas in clients and, therefore, inter-cell and intra-cell interference becomes an issue. Since the WiMAX community began to claim licensed spectrum for its technology, the debate has raged about the relative spectrum efficiencies of the TDD and FDD. Clearly, TDD offers the best flexibility in adjusting to the rapidly changing load factors of the IP world. This flexibility comes at the price of reduced spectral efficiency. Because the instants of uplink or downlink transmissions of a given device are unknown – or known only to the base station which a client is connected to, the spectral separation between cells has to be large, so that the required SIR at each client is maintained at all times. For truly high speed data transmission that means 20 dB or more. In a free space environment, this means at least three cell diameters separation between cells using the same frequencies. In a dense urban setting, this may drop to two cell diameters. The implication is that 16 respectively 9 frequency channels[35] would be needed to achieve the same data rates as an FDD system. On the other hand, FDD systems may waste some of their capacity if the offered load does not match their up/down capacity ratio. This is less the case in systems using OFDMA. Since both systems are subject to dynamic optimization, the actual performance differences between TDD and FDD may well prove to be minor.

The supreme irony of the WiMAX/LTE case is that as TDD WiMAX developed to improve its spectral efficiency in handling asynchronous data services, the mobile users moved towards much heavier use of video and streamed TV and those are served best with FDD systems. Notwithstanding their FDD heritage, early LTE designs were actually TDD systems that could compete with WiMAX – at least on paper.

[34] The actual difference in performance and spectral efficiency between these two modes of operation is subject to much debate. In general, the price paid for the improved flexibility offered by the TDD is a lower spectral efficiency.

[35] See also Chap.7, Sect. 7.2.4 *Frequency Sharing in the family*.

6.7.3 The Medium Access Protocol

IEEE 802.16 is a point-to-multiple-point wireless technology for the outdoor environment: one Base Station serves multiple subscriber stations. It is designed to accommodate Quality of Service (QoS) for latency-sensitive multimedia traffic. QoS considerations were an important aspect in the development of 802.16 from the very beginning. Four classes of services are defined: constant bit rate service, real-time variable bit rate service, non-real time variable bit rate service and best effort service. The 802.16 Physical Layer and Medium Access Control specifications have been designed to accommodate these special needs.

Down-stream signalling from the Base Station to the subscribers prepares up-stream from the subscribers to the Base Stations. This applies to established traffic flows, which are controlled by the scheduler of the base station. Requests for new traffic flows may be signaled in response to polling by the Base Station, as well as by subscribers contending for access to the Base Station during short "Contention Slots" provided by the protocol. In addition, the more recent versions of the standard provide an OFDMA capability, which significantly improves the capacity of the system, notably in the uplink (see Sect. 6.5).

From a spectrum sharing point of view, TDD/TDMA is the more interesting technology since it may see use in both license exempt spectrum and licensed spectrum that is adjacent to license exempt spectrum. In addition, there is overlap with wireless LANs in the 2.5 GHz band and in the 5.4–5.8 GHz band. In the former case, the issue is adjacent band coupling only; in the second case, all variations of coupling and interference may obtain.

6.7.4 Co-Existence with Other Technologies

The properties of the IEEE 802.16 architecture – notably its centralized control of medium access by all elements of a network – delimit the possibilities for co-existence with other systems. For FDD systems, intra-system coexistence is hardly an issue: proper planning of downlink and uplink frequencies assures minimal interference – as is true of conventional cellular systems.

The more interesting case is co-existence between TDD systems and, by extension, between TDD and FDD systems. The former easily generalizes to coexistence with any other non-IEEE 802.16 system. Various approaches have been considered for coordinated spectrum sharing; they all come down to separation in frequency or, if that is not possible, operation in alternate time windows with anther system. Clearly, that other system has to be a centrally controlled TDMA system as well, in order to be able to cooperate in this manner. Where this is not the case, a detailed case analysis is needed. The case of IEEE 802.16 and IEEE 802.11 technologies is investigated in more detail in Chap. 9, Sect. 9.2.

6.8 Summary

This chapter shows that spectrum sharing features as incorporated into some of the leading wireless technology standards are quite diverse as well as optimized for the system they are part of. With the exception of the optional adaptive hopping in Bluetooth and the optional energy sensing based LBT in the IEEE 802.11 standard, there are no examples of generic spectrum sharing capabilities. This reflects a choice in favor of maximizing spectral efficiency of a given technology, rather than offering such efficiency on the altar of supposedly fair and equitable use[36] of spectrum resources.

[36] See also Chap. 11, Sect. 12.2.5 *Fairness*.

Part II
The Practice of Spectrum Sharing

Part I provides a theoretical basis consisting of material about radio regulations together with an analysis of the factors that affect spectrum sharing and a discussion of the modes and mechanisms of spectrum sharing.

Part II provides investigations of actual commodity wireless technologies and evaluates the consequences of the design choices made. It does so using the IEEE 802.11 wireless LAN standard and its implementations as example – the reasons include that this technology has been in use for a long time, that it has seen a very wide adoption in the market place and that its properties and performance continue to evolve. Its basic approach to channel sharing – the listen-before-talk approach has become a model for other applications and technologies and its Dynamic Frequency Selection capability has become a prime example of spectrum sharing between primary and secondary spectrum users.

Chapter 7 provides an investigation of spectrum sharing among like systems. It uses wireless LAN technology as example of how small variations in technical details – defined by various Amendments of the IEEE 802.11 standard – have affected its sharing behavior.

Chapter 8 analyzes the effects of sharing among dissimilar commodity technologies – Bluetooth, Zigbee and Ultra Wide Band – using wireless LANs as a common element.

Chapter 9 addresses the important problem of how a commodity wireless system like a wireless LAN can share spectrum with an important primary spectrum user, such as weather radar systems. It provides a detailed analysis of this sharing regime, as well as considering sharing with fourth generation cellular system, sharing TV White Space Spectrum and Cognitive Radio techniques.

Chapter 10 investigates how radio resource management techniques can be used to optimize the performance of license exempt networks in the presence of interference. Wireless LANs are used as the example, because they represent the most advanced use of this kind of commodity technology.

Chapter 7
Spectrum Sharing with Wireless LANs

7.1 Introduction

This chapter looks into the sharing of spectrum within one family of commodity wireless technology: the wireless LAN family. Wireless LANs based on the IEEE 802.11 standard have become a de facto presence in everyday work and pleasure pursuits. Therefore, wireless LANs present an interesting case study of what can be achieved with spectrum sharing among similar systems.

Since its inception as a simple local access tool and private/home networking technology, in the early 1990, wireless LAN technology has evolved technically – e.g. the High Rate PHY (802.11n-MIMO) and it has moved into other areas of application – e.g. wireless sensors and Outdoor Mesh Networking. The latter type of use includes both Public Safety networks, industrial networks and consumer oriented public access networks. Challenges in the area of spectrum sharing occur in both the mixing of technologies and in the area of outdoor mesh networking.

The question is relevant of how the newer versions of this technology share the unlicensed spectrum with the older versions, which are based on the 802.11b/g and 802.11a standards. To understand the issue, one has to look at the basic elements of both the parent and the progeny technologies.

7.2 Legacy Wireless LANs

The behavior of wireless systems is largely determined by the properties of their transceivers and by their medium access protocol. This also applies to spectrum sharing. The following sections analyze the interaction of the evolution of the 802.11 transceiver specification and the medium access protocol. Also addressed is the effect of the latter on the throughput and channel capacity, achievable with a given transceiver.

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology, DOI 10.1007/978-94-007-1585-1_7, © Springer Science+Business Media B.V. 2011

7.2.1 The Early Versions: Spread Spectrum and CSMA/CA

The essential elements of the original IEEE 802.11 specification were a regulatory constraint called spread spectrum transmission and the invention of spectrum sensing combined with collision avoidance: CSMA/CA. The former was defined by the FCC as a means to prevent aggressive, winner-take-all technologies from becoming dominant users of the license exempt frequency bands they opened up in 1988: the 915 MHz band and the 2.4 GHz band – both designated as available for “Industrial, Scientific and Medical” applications. The spread spectrum techniques were prescribed: Frequency Hopping with a minimum of 79 hops and direct sequence spreading with a processing gain factor of at least 10.

CSMA – Carrier Sense Multiple Access was the mechanism developed by the industry to assure efficient channel sharing without the need for a central controller. See Sect. 6.2 for more details. CSMA/CA assumes a common view of the medium by all nodes participating in network i.e. “all-hear-all.” It makes use of carrier detection rather than energy detection: the decoding of the training sequence of other wireless LAN transmissions.[1] A hidden assumption in the initial design is that all nodes of the network use the same power level and have the same receiver sensitivity. Once this assumption no longer holds, spectrum sharing issues appear.

7.2.2 The Introduction of CCK Modulation

The importance of the *all-hear-all* rule came to light when the FCC agreed to a rule change that dropped the spread spectrum requirement. This made it possible to develop a much faster modulation scheme called Complementary Code Keying or CCK. This became part of the standard in 1999 as Amendment 802.11b. CCK offers maximum raw data rate of 11 Mb/s and uses the same bandwidth and media access method defined in the original standard. Therefore, one would expect no issues with spectrum sharing. However, progressing insight indicated that a shorter receiver training was possible – this reduced the training overhead from 192 to 96 μs and yielded a major throughput increase.

At 11 Mbps, a typical frame length of 234 bytes takes only 500-odd μs and, therefore, the short pre-amble increased the efficiency to 40%, much better than the 34% efficiency figure associated with the long pre-amble. See Table 7.1. Note that the figures reflect a fully loaded channel, not an individual device throughput.

The price paid for this efficiency benefit was loss of the *all-hear-all* property of the signaling: legacy IEEE802.11 devices could not decode the short pre-amble

[1] The alternative is simple energy sensing but this has two disadvantages: energy detection is far less accurate and it would cause wireless LANS to defer to other types of transmitters.

Table 7.1 IEEE 802.11b protocol efficiency, long preamble

Frame type	TCP overhead	Data avg	Media avg	File transfer
Data field size –	64	234	577	1,500
Slot duration	20	20	20	20
Avg contention res, slots				
Avg contention res, time	120	120	120	120
Pre-amble signaling rate (Mb/s)				
Pre-amble 1 length	144	144	144	144
Pre-amble 1, duration	72	72	72	72
PLCP header length	48	48	48	48
Pre-amble 2, duration	24	24	24	24
Data field rate and duration	47	170	420	1,091
SIFS	40	40	40	40
Ack pre-amble	72	72	72	72
Ack duration	3.6	3.6	3.6	3.6
Total TX OP time	378	502	751	1,423
Overhead, time, µs	332	332	332	332
Protocol efficiency	0.12	0.34	0.56	0.77

Table 7.2 IEEE 802.11b protocol efficiency, short preamble

Total TX OP time	306	430	679	1351
Overhead, time, µs	260	260	260	260
Protocol efficiency	0.15	0.40	0.62	0.81

and therefore would ride roughshod over distant[2] CCK transmissions – and vice versa (Table 7.2).

As a consequence, short pre-amble only operations had to wait until most legacy devices would have been replaced by the new "11b" devices. Thanks to the small installed base of the former at the time and the vast success of the latter, this did not take much time: from 2003 onwards practically all devices sold were either 802.11b or 802.11b/g.

7.2.3 RTS/CTS – The Partial Solution

The 1999 version of the "11b" standard includes the RTS/CTS mechanism; it is intended to reduce the so-called hidden node problem: propagation conditions may be such that a transmission is started by a device that is in range of the intended recipient, but not in range of the transmitter. RTS/CTS allows a transmitter/receiver pair to silence its neighbors with a Request to Send control frame and the Clear to Send response. A station is only allowed to send an RTS frame if none of its

[2]The IEEE 802.11 standard does specify an energy detect base channel check but the threshold for that is 20 dB higher than the sensitivity for the standard frame header.

neighbors are transmitting and the station has not heard a CTS covering the time instant when it wants to transmit. Both the RTS and CTS frame contain a NAV – the Network Allocation Vector that is used by receiving stations to set a virtual carrier detect duration. Notably, the CTS is assumed to take care of hidden node problems. Proper behavioral response to RTS and CTS control frames is required for all IEEE 802.11 devices.

RTS/CTS is far from ideal[3]: it adds overhead that grows disproportionally as the frame size goes down or the data rate goes up, it prevents transmissions that could have been successful and, worst of all, it is not very effective in preventing collisions.

7.2.4 The Introduction of OFDM

OFDM was introduced to further increase the throughput of the already dominant wireless LAN technology. The transition to OFDM was another event that showed the importance of the *all-hear-all* rule. The OFDM training sequence was very different from that for CCK or the spread spectrum modes; its length was reduced to 52 μs from the 90 μs of CCK. The combination of faster data transmission and the fixed channel training sequences caused the actual throughput of the higher OFDM rates be a lot less than the raw data rate would suggest. The reduction of the training sequence improved throughput a little, but not much. Tables 7.3 and 7.4 show the components of a complete unicast cycle for two data rates: 12 and 48 Mb/s.

They show that protocol efficiency rapidly drops with increasing data rate and decreasing frame size. The implication is that there is considerable margin that can be leveraged by applications and networking software by using larger frames. The trend towards heavier use of video, also in simple Internet applications, fits this context nicely.

The co-existence with legacy devices, including those using CCK, was initially[4] based on an *energy* detect level of −62 dBm in a 20 MHz channel.[5] This the older versions did not have and, therefore, newcomers were at a distinct disadvantage: in a mixed environment, the older versions would not defer for the OFDM devices, whereas the latter would – but only for nearby legacy devices. This was recognized early on and the Extended Rate Physical Layer Amendment of 2003 added that:

> the ERP modulations [..] have been designed to co-exist with existing [implementations]. This coexistence is achieved by several means, including virtual carrier sense (RTS/CTS or CTS-to-self), carrier sense and collision avoidance protocols, and MSDU fragmentation.

It is clear that it is left to the designer to devise ways to protect his implementation from collisions or pre-emption by other stations operating according to any of

[3] See Sobrinho et al. [116] and Shi et al. [112].

[4] See IEEE 802.11a-1999, clause 17.3.10.5.

[5] This was later generalized to include 5 and 10 MHz channels with pro rata settings for the ED threshold.

Table 7.3 IEEE 802.11a protocol efficiency for 12 Mb/s

	Frame type			
Parameters	TCP overhead	Data avg	Media avg	File transfer
Data field size	64	234	577	1,500
Slot duration	9	9	9	9
Avg contention res, slots				
Avg contention res, time	54	54	54	54
Pre-amble sign rate				
Pre-amble 1 length	120	120	120	120
Pre-amble 1, duration	20	20	20	20
PLCP header length	0	0	0	0
Pre-amble 2, duration	0	0	0	0
Data field rate and duration	43	156	385	1,000
SIFS	18	18	18	18
Ack pre-amble	20	20	20	20
Ack duration	3.3	3.3	3.3	3.3
Total TX OP time	158	271	500	1,115
Overhead, time, µs	115	115	115	115
Protocol efficiency	0.27	0.57	0.77	0.90

Table 7.4 IEEE 801.11a, protocol efficiency for 48 Mb/s

Total TX OP time	124	152	209	363
Overhead, time, µs	113	113	113	113
Protocol efficiency	0.09	0.26	0.46	0.69

the "legal" 802.11 operating modes. The CTS-to-Self is a mechanism that allows a transmitter to silence other nearby transmitters by sending out a "response" to a virtual Request-to-Send. The effect of a CTS-to-Self is that the stations around a transmitter are forced to delay their actions until the NAV of the CTS-to-Self has expired. Like the RTS/CTS mechanism, it is not fully effective and it reduces channel capacity by blocking transmitters whose transmissions would not affect the transmissions of sender of the CTS-to-Self. In practice, all of this complexity proved a bit of overkill: the rapidly growing market absorbed new versions of the OFDM technology without much impact on existing installations.

7.2.5 Frequency Re-use in the Family

The above story of the evolution of IEEE 802.11 does not address spectrum re-use. Although the CSMA/CA mechanism prevents most collisions under many circumstances, it does not prevent all of them. As noted above, the carrier defer threshold for the CA mechanism lies well above the minimum useful signal level: in the case of OFDM, it is −82 dBm for a 20 MHz channel or 19 dB above the noise floor. For energy detection, the difference is 39 dB. In the conventional access point based deployments, this large margin need not be a serious issue: it increases the area in

which clients are exposed to the risk of an SIR than supports only lower transmission rates. In large scale networks, the use of carefully selected channel plans.[6] However, the picture changes with the introduction of Direct Link protocols, which allow the role of the access point to be much reduced or even eliminated. Direct Link protocols increase the opportunities for failure of the CSMA/CA mechanism.

Spectrum sharing among family members tends to be a more benign affair than spectrum sharing among dissimilar systems. As described in the preceding section, even though they are all part of the same technology family, the family members do not share spectrum on an "equal basis" among each other. This is the unavoidable price of progress: new designs make the older ones obsolete or less desirable *because* full backwards compatibility would prevent the new designs from achieving their full performance.

7.2.6 High Rates Mean High SIR Margins

The new high levels of performance come at the price of larger SIR values and, therefore, shorter operating range. However, the energy generated by the new types of transmitter remains at the same level and so does their interference potential. The implication is a significant reduction in the ratio between the useful coverage area of a transmitter and the collision area that energy creates for other high rate systems. This ratio, which is known as the frequency re-use factor, is rooted in the SIR of the systems involved. It is a major element in determining the net capacity in bits/hertz/area that a given technology can achieve. The table below is based on a 17 dBm IEEE 802.11a transmitter operating at 5.6 GHz and a pathloss model that includes free space up to 8 m and −10 dB/octave beyond that. The receiver is assumed to have a bandwidth of 20 MHz. That gives a noise floor of −101 dBm; the noise figure is assumed to be 6 dB. The SNR required is assumed to be 7 dB for OFDM BPSK, coding 1/2. For 64QAM, 2/3 coding, SNR = 21 dB is used. Fading is not included in these figures. In order for a wireless LAN device to suffer very little loss of range, an interfering signal must be less than −101 dBm.

7.2.7 Effective Coverage Area and Channel Capacity

The effective coverage area is given by the square of the relative range figures. The number of channels needed to completely cover a given area with full capacity is given by the square of the re-use distance ratio. This is the ratio between the operating range of the victim and the sum of that range and the interference distance. For 48 Mb/s (64QAM with 2/3 coding) the operating range works out at 50 m, the

[6] See Chap. 10.

Table 7.5 IEEE 802.11a, large scale throughput figures

	BPSK-1/2	QPSK-1/2	16QAM-1/2	64QAM-2/3
Data rate (Mb/s)	6	12	24	48
SNR (dB)	7	10	15	23
Coverage distance (m)	133.5	108.3	76.4	50.3
Protection distance (m)	312.1	312.1	312.1	312.1
Re-use distance (m)	445.6	420.4	388.5	362.4
Raw Throughput @ 577 B/frame (in Mbp/s)	4.97	8.55	13.34	18.54
Re-use distance ratio	3.34	3.88	5.09	7.2
Net channel capacity in Mb/s	0.45	0.57	0.52	0.36
Spectrum capacity (over 19 channels) in Mb/s	8.48	10.78	9.80	6.78

Table 7.6 EEE 802.11a, isolated cell, throughput figures

Net channel capacity in Mb/s	3.80	6.17	8.53	9.72
Spectrum capacity (over 19 channels) in Mb/s	72.28	117.27	162.05	184.69

interference distance at 312 m. The resulting distance ratio is 7.2, the required number of channels would be 52. The number of available channels in the 5 GHz band is only 19 and therefore the maximum throughput per channel is $19/52 \times 18.54$ Mb/s = 6.77 Mb/s.

This figure is scale independent – from a few meters between Access Points up to the practical maximum distance. At the range calculated above, the highest data rate would deliver 6.78 Mb/s over an area of 7,850 m^2 or 863 bps/m^2. At an Access Point spacing of 10 m, that figure jumps to 216 kbps/m^2. In practice, the difference is not as large: contention resolution and pre-amble overhead reduce the differences in net throughput at all bitrates (see Table 7.5).

The different data rates give roughly the same large scale capacity figures – because of two factors: the fixed overhead per frame and the increase in SIR with higher bitrates. Both take a higher toll at higher data rates. This picture changes if the pathloss between the "cells" is increased, e.g. by figures that correspond to a typical suburban situation: room size of ~20 m^2, wall attenuation of 13 dB and spacing between dwellings of 16 m. This gives a total attenuation of 13 + 12[7] + 13 = 40 dB. This largely eliminates interference and in that case the above channel capacity differences become more representative of the underlying difference in bit rate (see Table 7.6).

These latter figures dominate the use of wireless LANs in isolated homes and well insulated apartment buildings. In offices and low-cost high-density housing, the figures of Table 7.5 are more relevant. However, both are extremes and practice may prove more benign in many cases. This expectation is further supported by the fact that the IEEE 802.11 protocol gracefully adapts to impairment of propagation conditions.

[7]Estimated attenuation of a 5 GHz signal through shrubs and trees over a distance of 16 m.

7.3 MIMO Wireless LANs

MIMO – multiple input/multiple output – is a technique that allows multiple independent transmissions between two devices over separate transmit/receive antenna sets. MIMO leverages the spatial diversity of the physical radio channel: e.g. four such sets can deliver four times the data rate.

When MIMO was added in 2007 (to be formally approved in 2009), another dimension of co-existence issues was introduced: both the spatial multiplexing and 40 MHz channel options introduced many differences that required special provisions in the standard[8] in the form of three modes or physical layer frame formats:

- The mandatory Non-HT format, which includes the OFDM and ERP frame formats; this format provides full interoperability with "pre-MIMO" implementations.
- The mandatory HT-Mixed format, which includes pre-ambles which can be decoded by pre-MIMO implementations; this format allows the co-existence as equals of pre-MIMO and MIMO implementations – provided the 40 MHz option is not used by the latter.
- The optional HT-Greenfield format, which allows for a shorter, more efficient MIMO pre-amble. Decoding HT-Greenfield pre-ambles is mandatory of all HT devices. This assures co-existence as equals of HT-Greenfield and non-HT-Greenfield implementations.

Clearly, from a new installation point of view, the Greenfield option is attractive – it offers the best performance, but it has a price – all wireless LAN devices must be capable of running in HT-Greenfield mode. If a mixed population of clients must be supported, fall back to Non-HT mode is required. The HT-mixed mode allows logically separate populations of HT and legacy devices to co-exist, but not interoperate.

7.3.1 MIMO Pre-ambles and Protocol Efficiency

The IEEE802.11n Amendment specifies a number of pre-ambles to match the three modes of operation described above.

- The Non-HT mode uses the 802.11a pre-amble of 20 μs.
- The Mixed-HT mode uses the 802.11a pre-amble of 20 μs plus 12 μs HT signaling and n × 4 μs per spatial stream. For a two-stream system, this adds up to 40 μs.
- The Greenfield-HT se mode uses 24 μs plus (n − 1) × 4 μs for n spatial streams. This gives 28 μs for a two-stream system.

This gives the following protocol efficiency figures for 13, 52 and 104 Mb/s systems using two streams (Table 7.7).

[8] See IEEE 802.11 2007, Clause 19.

Table 7.7 Examples of protocol efficiency for MIMO wireless LANs

		Frame type			
Mode	MSC type-data rate	TCP overhead	Data avg	Media avg	File transfer
Non-HT	MCS 8-13 Mb/s	0.25	0.56	0.76	0.89
	MCS11-52 Mb/s	0.08	0.24	0.44	0.67
Mixed	MCS 8-13 Mb/s	0.20	0.48	0.70	0.86
	MCS11-52 Mb/s	0.06	0.19	0.37	0.60
	MCS13-104 Mb/s	0.03	0.11	0.23	0.43
Greenfield	MCS 8-13 Mb/s	0.23	0.52	0.73	0.88
	MCS11-52 Mb/s	0.07	0.22	0.41	0.64
	MCS13-104 Mb/s	0.04	0.12	0.26	0.47

The differences in protocol efficiency between these modes are minimal – the picture is dominated by the frame size and by the data rate. The 802.11 Working Group understood the benefit of long frames and provided a means to achieve that: concatenation of MAC layer frames into larger transmission units. This done either by chaining frames together and using block acknowledgement or by combining frames – this is called frame aggregation. In the latter case, MPDUs are combined together into a single PHY service data unit that is transmitted as one frame with a single pre-amble. This allows frames up to 65,535 bytes. The jury is still out on how well such large frames survive the transmission process: at 162 Mb/s (3 × 3 MIMO, 40 MHz, 16QAM, coding rate =1/2) transmission time is 3,288 μs (3,240 μs for the data and 48 μs for the pre-amble). The typical channel is stable only during a few milliseconds. This limits the achievable frame size/data rate combinations that yield a short transmission time, e.g. less than 1 ms. In case of the example cited, this would be ~20,000 bytes. To keep the frame error rate to <10% requires a bit error rate of $\sim 0.5 \times 10^{-6}$. In practice, this requires a very large SNIR which is not likely achievable for very large, high density installations. Very high SNIR values degrade the spectrum re-use factor resulting in performance loss – which counteracts the benefits of the higher protocol efficiency. Further, the impact of small channel disturbances or processing errors on the frame error rate of such large frames is not clear. These considerations suggest that a conservative frame size of 2,000 bytes should be used as maximum for network planning and performance estimation purposes. At such a frame size, the benefits of the higher rates are small indeed.

7.3.1.1 Mixed 20/40 MHz Channels

The IEEE 802.11 standard provides for optional 40 MHz channels. This 40 MHz channel option not only provides choices for wireless LAN implementers and users, it also introduces the need to avoid conflicts in spectrum sharing with legacy equipment. For 20 MHz operations, no special arrangements are required – other than setting the HT mode of operation to "20" instead of "20/40".

Forty megahertz operations are more complex to put into effect because the spectrum sharing conditions are more complex: mixed 20/40 MHz operation may be

needed because not all devices may be able to operate in 40 MHz mode and/or that a second channel is not always available. The initial operating channel is called the primary channel, the other channel is known as the secondary channel. The secondary channel may be above or below the primary channel.

The standard specifies a number of means to manage the 20/40 sharing conditions:

1. Least used channel selection at start up
2. Using CTS-to-Self to avoid interference[9] from legacy wireless LAN devices
3. Signaling 40 MHz intolerance
4. Independent clear channel assessment on both primary and secondary channel
5. Changing channel or channel width if an overlapping BSS condition arises or disappears
6. Special operating modes such as Phased Co-existence Operation (see below)

Items 4–6 of this list are of special interest from a spectrum sharing point of view. Clear Channel Assessment (CCA) determines if and when a device will transmit. Although the rules of CCA for 40 MHz operation in the standard are complex in details, the principle is simple: both channels must be monitored for signals; the threshold for HT frame detection is −82 dBm, and for energy detect it is −62 dBm. Normally, only energy detect is used in the secondary channel – except when any 40 MHz signal is received. Note that the implication is that 20 MHz transmissions on the secondary channel are ignored in the CCA procedure until it is time to transmit. The rule for starting a 40 MHz transmission is a down-counted contention window plus a clear condition on the primary channel AND a clear condition on the secondary channel for the duration of a PIFS. Since that PIFS period is very short, it could fall within a sensing period of a contending station. If this is the case, the other station is pre-empted and thus this PIFS check confers an advantage on the dual channel transmitter.

In addition, the standard provides a detailed procedure for "overlapping BSS scanning" on the secondary channel; the following captures the essentials. The scanning must be done on the "to be" secondary and it applies only to devices that are active more than 0.25% of the time, the scan interval has to be less than 300 s and the scan dwell time during that interval has to add up to ~200 ms, the minimum dwell time being 20 ms. This comes down to checking for 20 ms every 30 s, i.e. not very frequently. The underlying assumption may well be that BSS'es do not appear and disappear frequently.

In "Phased Co-existence Operation," an Access Point which operates in 20 MHz has to announce when it wants to perform a 40 MHz transmission. On the next clear channel condition (on both channels, see above), the Access Point will send out a CTS-to-Self on both primary and secondary channels to silence its neighbors before initiating a 40 MHz transmission. Note that this does assure that the secondary channel is free at the receiver. Upon completion of its transmission, the access point will revert back to 20 MHz mode.

---

[9] This is not always effective – see Sect. 7.2.3

7.3.2 Frequency Re-use Considerations

MIMO technology has been heralded as the breakthrough towards ever faster data rates for wireless LANs. That perspective is seductive and the rapid adoption of MIMO/11n technology in the market place speaks to that seductiveness. However, none of the analyses of MIMO wireless LANs performance pays much attention to spectrum re-use limitations.

The available SNIR determines the required separation distances for frequency re-use. Below, we make use of the same model and parameters as used in Sect. 7.2.4: looking at the throughput of a single device – e.g. an Access Point – taking into account that its frequency of operation is being (re-)used nearby by other devices, possibly belonging to a separate "network." MIMO requires more complex calculations and, given that digital sampling and digital processing are subject to limitations in precision, a better SNR is needed than for single stream OFDM processing. Adding 4 dB to SNR values used for calculation legacy efficiency gives the following results for large-scale and isolated deployments of MIMO wireless LANs for the multi-media frame type and the Mixed HT mode of operation using 2 × 2 MIMO (Table 7.8).

Clearly, there is little benefit to be had from increasing the data rate beyond 52 Mb/s in large scale, high density deployments and this applies regardless of frame size. The spectrum re-use factor – shown in *italics* – is increasingly unfavorable for higher rates: the number of channels needed for full throughput is respectively 18, 25 and 45. With only 19 channels available, data rates above 52 Mb/s are subject to frequency re-use limitations: almost a factor 3 in the case of the 104 Mb/s rate. This means that at 104 Mb/s only one-third of the full throughput can be realized. This applies even more strongly for the higher data rates.

In other words, the higher data rates only make sense in isolated deployments – in homes, small businesses, etc. This extends to 40 MHz operations: although the required SNIR drops when a 40 MHz channel is used for the same data rate, the gain is significant at the lower speeds, but less so at higher speeds. For consistency reasons, the table below uses twice the rates used in Table 7.9, rather than the real rates (27,108 and 216 Mb/s) that apply for 40 MHz channels.

If one ignores that 40 MHz channels give a data rate that is about 4% higher than twice the 20 MHz rate, 2 × 2-40 MHz and 4 × 4-20 MHz deliver the same throughput but in half the spectrum. The main difference between these two ways of achieving the same throughput is the cost of the implementation.

Table 7.8 Examples of net channel capacity of MIMO wireless LANs in 20 MHz channels

		Spectral efficiency in Mb/s		
HT = 2 × 2 MIMO SDM system		13 Mb/s	52 Mb/s	104 Mb/s
Large scale deployment	Legacy	0.57	0.52	
	HT	0.80	1.18	2.73
Isolated deployment	Legacy	6.17	19.29	
	HT	7.41	14.00	49.57

Table 7.9 Examples of net channel capacity of MIMO wireless LANs in 4 × 4 and 40 MHz channels

MIMO SDM system		Spectral efficiency and *re-use factor*			
		13 Mb/s	26 Mb/s	104 Mb/s	208 Mb/s
Large scale deployment	2 × 2 HT-20	0.80	1.18	2.73	
		3.48	*4.05*	*5.33*	
	2 × 2 HT-40 4 × 4 HT-20		1.60	4.72	5.46
			3.48	*4.05*	*5.33*
Isolated deployment	2 × 2 HT-20	7.41	14.00	49.57	
		1.14	*1.18*	*1.25*	
	2 × 2 HT-40 4 × 4 HT-20		14.82	55.98	99.13
			1.14	*1.18*	*1.25*

7.3.3 Transmit Beamforming

The transmit beamforming capability defined by the IEEE 802.11n standard is optional. The spatial resolution of the beamforming increases with the number of antennas – in the case of four transmit antennas, the number of discrete beam positions is 12. Compared to straight spatial multiplexing, it delivers an SNR improvement of 4 db for a 2 × 2 system to 5 dB for 4 × 4 system at the lowest bitrates. This gain drops to zero at the highest bitrates. However, in case of 4 × 2 systems, SNR gain can be as high as 11 dB for the highest bit rate –relative to a straight 4 × 2 system. The benefit of beamforming lies in enabling the use of higher bitrates in certain conditions and device configurations.[10] Provided the beamforming transmitter has sounded the channel, MCS differentiation per antenna can be used to optimize the throughput. Beamforming is useful in asymmetrical systems in which one device has four antennas and the other two. Transmit beamforming is then used by the four antenna device whereas the return path uses 2 × 4 MRC.

From a spectrum sharing point of view, transmit beamforming has the negative property that complements its positive property of delivering energy selectively in a certain direction: less energy will be delivered in other directions than the preferred one. This affects the CSMA/CA mechanism by effectively "hiding" the beamforming device from many devices in its surroundings.[11] The impact increases with the number of antennas used by the beam former. Therefore, the receiver of the "beam formed" transmission is at risk of interference from devices that did not detect such a transmission. Depending on the density of devices, this interference may well nullify the gains of beamforming. This suggests that beamforming works best in cases where it is least needed: isolated wireless LAN networks that are not subject to significant co-channel interference.

[10] See Perahi and Stacey [101].

[11] See also Sect. 6.4.2.4.

7.4 Outdoor Wireless LAN Mesh Networks

The early years of the twenty-first century saw high hopes and audacious business plans for free, ubiquitous wireless LANs everywhere. City councils approved funds for helping to build such networks, and pundits praised the coming revolution of "free broadband everywhere," based on mesh networks built from Access Points using "Wi-Fi" technology. The promise of ultra high data rates, made possible by MIMO technology, brought some to think that wireless LAN meshes could replace wired infrastructure. Reality soon caught up with many of these plans and dreams, but the fact remains that mesh networks are very useful in many environments and applications – ranging from a squad of police cars to a city-wide tourist information network. The IEEE 802.11 Working Group recognized this and developed a "Mesh" extension of its standard, aimed at serving a variety of applications: from fixed networks to handheld mobiles. In this section, we look at the particular aspects that set apart outdoor wireless LANs from their indoor counterparts and that bear on spectrum sharing issues. The focus is on the mesh network itself rather than the wireless LAN clients it serves.

The largest difference between indoor and outdoor operating conditions of wireless LANs are not the extremes of humidity and temperature, but a highly unpredictable propagation environment. Indoor environments show large attenuation factors beyond a few meters from the transmitting antenna – caused by obstacles, human beings, and walls. Although large signal strength differences may occur at the scale of half a wavelength, large areas of deep shadows tend to be absent due to reflections off objects and walls. A simple indoor pathloss model consists of two or three sections, as shown in Table 7.10. The actual distance varies with the general environment, e.g.: office spaces tend to be larger than homes. In some cases, the third segment can be ignored because signal strength beyond Segment 2 is low enough for it to be discounted.

Like most pathloss models, these simple multi-segment pathloss models are statistical in nature: they represent an average over many instances in time and direction.

Although the general "model" of outdoor propagation loss may be similar to the one used for indoor propagation, the typical outdoor model is less easy to determine: it varies with terrain, with building height and density, foliage density, humidity, etc. A major factor is the antenna height of the Access Point: the higher it is located, the longer the free space segment and the medium attenuation segments and the shorter the last segment in which signal blocking by buildings dominates. The cellular communications industry has built up vast experience on this subject and a variety of outdoor propagation models exist. However, these models fit the high

Table 7.10 A typical indoor pathloss model

Segment 1	Seg ment 2	Segment 3
6 dB/octave	10 dB/octave	20 dB/octave
First few meters	Short distances (<20 m)	Beyond outer walls

base station – low mobile configuration. Mesh networks consist of peer nodes located at low to medium heights. For this type of communication, few pathloss models exist. Table 7.11 gives some examples.

7.5 Hidden Nodes and Other Asymmetries

Table 7.11 does not capture the local diversity of the low-level outdoor environment: the city landscape may give Category 1 like propagation along a street between tall buildings, at right angles to that street, Category 3 conditions are likely to dominate. This diversity has serious consequences for the CSMA/CA protocol: the *all-hear-all* principle on which it is based is not likely to apply for the majority of nodes in a mesh network. However, the same conditions, albeit on a smaller scale, occur in many indoor environments as well. The effects include hidden node cases in which two nodes communicate, but one of them suffers interference from a third node which does not receive signals from one of the other two. An example is given in Fig. 7.1.

Table 7.11 Example of three categories of low level outdoor pathloss models

Environment	Segment 1 6 dB/octave	Segment 2 10 dB/octave	Segment 3 >20 dB/octave
Category 1: Flat rural	<1,000 m	1,000–5,000 m	>5,000 m
Category 2: Wooded rural/suburban	< 100 m	100–500 m	> 500 m
Category 3: Central city/high rise	< 50 m	50–100 m	> 100 m

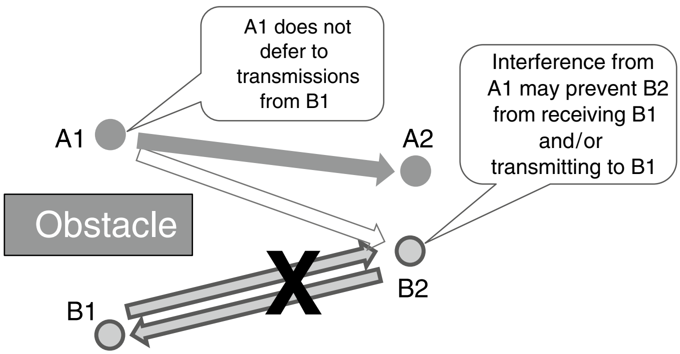

Fig. 7.1 Example of a hidden node condition

These and other asymmetries in propagation paths have a major impact on the medium access protocols by creating de facto "deafness" cases.

7.6 The Interaction of RF and the CSMA/CA MAC Protocol

The CSMA/CA protocol assumes that all devices – four in the above example – hear each other, but in practice this may not be the case: propagation conditions vary, transmission paths can be blocked and interference can cause a receive operation to fail. The exponential back-off procedure of IEEE802.11 is intended to take care of this. Unsuccessful transmissions require that the sender increases its contention window so as to prevent it from causing long-term damage to the network as a whole. Because the increase must be exponential, it can easily lead to a "winner take all" situation in which a few transmission failures lead to a catastrophic reduction in throughput.[12]

Another mechanism in the IEEE 802.11 protocol is the RTS/CTS exchange and its derivative the CTS-to-Self. The former assumes that the hidden node may be in range of either the sender (of the RTS) or the receiver (sending the CTS). The latter will silence only those in detection range of the transmitter, but not any hidden nodes. These mechanisms are marginally effective in alleviating the effects of propagation asymmetries.

Much work on this subject has been done by Rice University.[13] They showed that all wireless LAN mesh interactions can be analyzed based on a simple two-flow model[14]: Transmitter A talks to receiver "a" and Transmitter B talks to interactions between the devices (Fig. 7.2).

Any interaction between four or more nodes can be reduced to this model of four interacting elements. There are two primary interactions and four secondary ones. This gives a total of 12 unique interaction patterns. CSMA/CA is designed for the generic case in which all hear all. Blockage of some of the interaction paths affects the overall performance of the system – in some cases very heavily. There are four variants in which the two senders are invisible to each other. Conditions such as these occur in indoor environments, such as offices and factories, as well as in outdoor mesh networks (Fig. 7.3).

The effect of symmetrical incomplete state information is that the transmitters start transmissions more or less independently: whenever both attempt to access the channel, the one who is first forces the other into back-off. That increases the latter's contention window and effectively reduces its "aggressiveness" in acquiring the channel, to the benefit of the other sender. In practice, such a situation can persist for

[12] See also Sect. 6.2.1

[13] Rice University runs a mesh network called "Technology For All" in Houston that has proved to be a fruitful research tool.

[14] See Garetto et al. [52].

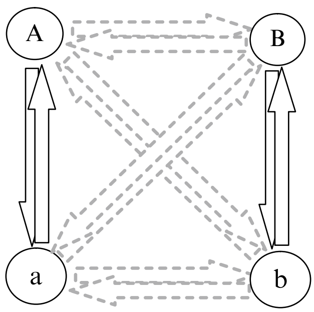

Fig. 7.2 General two-flow model

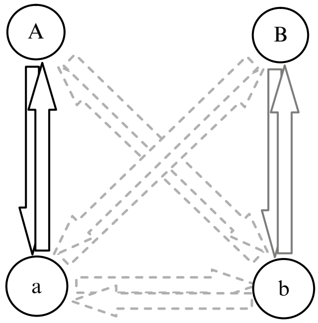

Fig. 7.3 Symmetrical incomplete state information variants

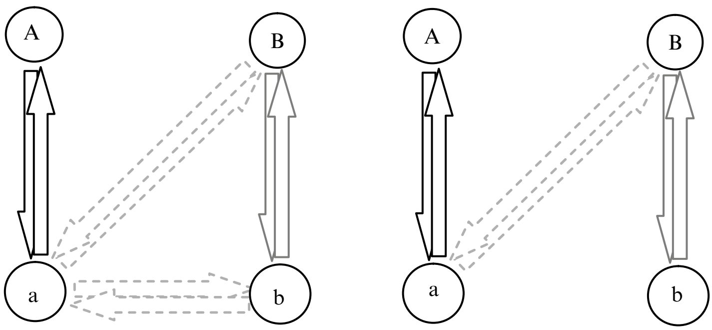

Fig. 7.4 Asymmetrical incomplete state information variants of two-flow interactions

some time; but eventually, the roles will switch. Therefore, the short-term unfairness is complemented by a long-term fairness (Fig. 7.4).

A more damaging effect is caused by asymmetrical incomplete state information: A will transmit at any time and the response from "a" will interfere with either B or "b" or both so as to drive them into back-off. Since there is no feedback to A, this situation will persist while A has traffic to send.

These complex situations cannot be addressed by simple "if this, then do this" rules: conditions may differ per network node and they may change with time. Therefore, dynamic adaptation along the lines discussed in Sect. 6.3 may well be needed. It promises to be a rich field for research and creative solutions.

7.7 Summary

As the preceding sections show, wireless LAN operations suffer from a number of inefficiencies that are due to a variety of reasons. These include different frame header formats and different channel widths, which lead to incompatibilities that can be solved only by means that require channel time, such as dual purpose headers and channel "negotiation" rules. Different modulation schemes lead to incompatibilities that cannot be solved, only ameliorated, e.g. through energy detection. This reduces the efficiency of the collision avoidance mechanism and increases the potential for asymmetrical coupling. Even if all devices in a given area operate with the same channel width, modulation scheme and header formats, the propagation conditions may cause pathological interactions between nodes that cause deafness and/or lead to starvation or start/stop throughput effects. The impact of all of the above depends on the modulation scheme: higher order modulations require a larger SIR and, therefore, they are more sensitive to propagation imperfections. The larger SIR also leads to a rapid drop in throughput with increasing bit rates for all nodes that share a given channel.

That wireless LANs have become the success story they are today owes much to the fact that the capacity they offer typically exceeds the user's demand. In that case, the above imperfections have no easily noticeable impact. This applies in most if not all households and offices. However, this will change if the trend towards replacing cables by wireless continues apace with the increasing demand for video applications over wireless. As the density of wireless LAN deployments increases, more robust modes of operation – i.e. lower SIR values – will be required. That means operating at data rates that are robust rather than "fast."

Chapter 8
Spectrum Sharing with Other Commodity Technologies

This chapter focuses on spectrum sharing between different wireless technologies. The examples are concerned with wireless LANs and other commodity wireless technologies such as Bluetooth, Zigbee and Ultra-Wide-Band. These are low power, license exempt devices, which are used in many applications and in various frequency bands. None of these technologies has been designed to share spectrum efficiently with other technologies. The results are instructive in that they show the sharing between dissimilar systems is not very efficient.

8.1 Coexistence of Wireless LANs and Bluetooth

8.1.1 Introduction

Bluetooth and wireless LANs are frequently present in the same devices such as notebooks and smartphones. Both are frequently used for "personal connectivity" and may be used simultaneously regardless of whether they are part of the same device or not. Both systems operate in the 2.4 GHz band, but the medium access approach is completely different: one relies on listen-before-talk, the other on statistical separation. Whereas an 802.11 device could in theory detect Bluetooth and, therefore, defer channel access to avoid colliding with it, the reverse is not true: Bluetooth does not use any form of channel sensing. Although this would suggest that 802.11 devices typically are the victim of nearby Bluetooth transmitters, practice proves more positive and more complicated. This is addressed in detail in the following sections.

The issue of sharing spectrum with "802.11" was recognized early on, but it took until the early years of this century for a formal publication on the matter.[1] The

[1] See IEEE 802.15.2-2003-Part 15.2: Coexistence of Wireless Personal Area Networks with Other Wireless Devices Operating in Unlicensed Frequency Bands.

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_8, © Springer Science+Business Media B.V. 2011

publication pointed towards a possible solution: some form of tight coupling that would allow both systems a fair degree of autonomy while reducing the mutual interference. The even more effective solution of "moving" 802.11 based systems to the 5 GHz band never saw much traction – until MIMO technology was introduced.

Although neither wireless LAN nor Bluetooth devices were originally designed with the existence of each other under consideration, the communication protocols for both types of devices are often robust enough and include mechanisms for error checking and correcting, as well as requesting that corrupted frames be resent. Research[2] has confirmed this. 802.11 devices may slow down their transmission rate and try again. Bluetooth devices hop away and try again when frames are lost. As a result, the impact of increasing levels of interference is at least a slowing of the data rate. In general, wireless LANs operate reliably in the presence of significant Bluetooth interference. Only under extreme conditions is it likely that wireless LAN communications will cease altogether because of such interference.

Nonetheless, the industry has pursued solutions. Adaptive Frequency Hopping (AFH) for Bluetooth devices is a representative solution which reduces interference by choosing hop frequencies that do not appear to be used by other devices such as wireless LANs. The implementation of such mechanisms greatly reduces the occurrence of interference between the two systems – at least in isolated conditions. However, the increasing use of smartphones and pads will likely lead to more interference between these and Bluetooth devices – if only because both have both 802.11 and Bluetooth capabilities. The fundamental questions are (a) how severe the interference will be, and (b) how it will impact the performance of wireless LAN devices and the Bluetooth devices. In order to answer these questions, one has to look at the static link budgets, as well as at the dynamic interactions.

Details of wireless LAN operation, including the carrier sense mechanism, is described in Sect. 6.2. Bluetooth operation, including its frequency hopping, is described in Sect. 6.6.

8.1.2 Link Budget Analysis

From a spectrum sharing viewpoint, wireless LANs and Bluetooth present an interesting case, not in the least because of the broader context. Both where designed roughly at the same time and both paid much attention to sharing of spectrum – but only with its own kind.

Because the modulation of Bluetooth differs from that of the wireless LAN, the latter has to rely on energy sensing to detect Bluetooth.

For CCK modulation (802.11b) the energy detect threshold is −80 dBm for TX power > 100 mW, −76 dBm for 50 mW < TX power and −70 dBm for TX power of <50 mW.[3] For the more widely used OFDM modulation, the recommended energy

[2] See e.g. Gerrior and Woodings [56].

[3] See IEEE 802.11-2007, clause 15.4.8.5.

detect threshold is −62 dBm for a 20 MHz channel width and a power level of 50 mW[4]: this is 20 dB above the minimum signal level for the lowest data rate. The difference between the energy detect thresholds for CCK and OFDM are worth noting: the DSSS specification protects others relative to the own transmitted power level, whereas the OFDM threshold only considers a minimum SIR margin needed for the highest data rate in a given channel width. For a 16.3 MHz channel, this threshold is 14 dB higher than the margin afforded by the CCK specification.

The consequences are that IEEE 802.11 devices typically ignore other types of devices unless these are operating nearby: 20 dB difference in signal strength equates to a difference in range of a factor 4–8, depending on the type of environment. For the typical Bluetooth device (Class 2) which operates at 0 dBm output power, the difference is smaller: the difference of 17 dB in power is largely compensated by the difference in bandwidth, again depending on the environment. The typical MUS level for an wireless LAN operating at 11 Mb/s it is −76 dBm, for 54 Mb/s it is −62 dBm; for Bluetooth it is roughly −70 dBm. Because wireless LANs adjust their data rate to local conditions, one can assume that the average MUS level for a wireless LAN device is −68 dBm. Using a propagation model with two exponents and a breakpoint at 8 m (2/8; 3.3/infinity), the interference distance works out at approximately 30 m.

Given these distances, spectrum sharing between wireless LANs and Bluetooth is no issue in practice. However, things change if these distances get a lot shorter – as is the case when having Bluetooth and wireless LAN transceivers operate on the same desk or having them located in and controlled by the same host device. Those scenarios require a dynamic analysis that takes into account the air interface behavior of both systems.

8.1.3 Protocol Interaction Analysis

Typically, Bluetooth is used as a personal area network – people use it to connect gear they use in close proximity to themselves. Since PC notebooks and PC desktops are also used as personal devices, the most frequently occurring interaction between wireless LAN gear and Bluetooth gear will be between wireless LAN clients and Bluetooth piconets. The latter will usually consist of a few devices that are controlled by a master device such as a PDA or a PC. The distances involved are short – a few meters – and therefore the distance factor is not taken into account in analyzing the interference.

The key to the interference assessment lies in the traffic patterns of the two types of device. Differences in the traffic patterns can create degrees of freedom that would not occur if both followed the same deterministic pattern. On the other hand, the interaction between deterministic and semi-random patterns does lead to

[4]See IEEE 802.11-2007, clause 17.3.10.5.

interference at the message level that can be quite destructive if the destroyed data is *control data* that governs device behavior.

8.1.3.1 Wireless LAN Traffic Patterns

Wireless LANs have two transmission types: broadcasts and unicasts; both are preceded by contention resolution.[5] The broadcasts are used for relatively short beacons and for the even shorter "here I am" keep alive messages generated by the IP layer. The intervals of the latter vary, the beacons are assumed to be sent at regular intervals e.g. every 100 ms. The unicasts carry data traffic – here the frame size ranges for a low of 40 – odd bytes to a 1,000 bytes. Because the beacons are sent at the lowest rate – i.e. 6 Mb/s in case of an OFDM system – the actual airtime may be larger than that of the data messages sent at full rate. Further, the unicasts are acknowledged – if correctly received – at the same data rate as the received unicast. However, the transmission of the ACK is not preceded by a check for a clear channel.

8.1.3.2 Bluetooth Traffic Patterns

Bluetooth tends to be used for peripheral functions like headset, mouse, and keyboard. Notably, headsets are used in "connection" mode, rather than in bursty modes like mouse and keyboard. A Bluetooth voice link uses the SCO service, in which both master and client are active continuously.

Unlike the wireless LAN, Bluetooth is a frequency hopper. The slot duration is equal to the hop interval: 1/1,600 s and within a hop, both downlink and uplink transmissions take place. Since there are 79 hop positions in the hop frequency table, the same frequency will be used 79/1,600 or 20 times per second on average. This is equal to ~50 ms intervals.[6] Since the wireless LAN channel is roughly 20 times the width of a Bluetooth hop, the "hit rate" at which interference can occur is 50/20 = every 2.5 ms or 400 times per second. The Bluetooth hop dwell time is .625 ms, the transmission part half of that.

8.1.3.3 Wireless LAN – Bluetooth Interaction Patterns

The transmission times of wireless LAN frames vary with bit rate and frame length. The OFDM training sequence and rate signaling takes 20 μs; that together with a 2,000 byte frame at 24 Mb/s gives 20 + 666.7 = 686.7 μs, whereas a beacon frame of 64 bytes at 6 Mb/s takes about 105.3 μs.

Given the difference in transmit power levels, for any nearby wireless LAN receiver, the transmissions of a nearby Bluetooth device are the more damaging.

[5] See Sect. 6.2.

[6] Due to the random nature of the hop sequence, the actual interval may be much shorter.

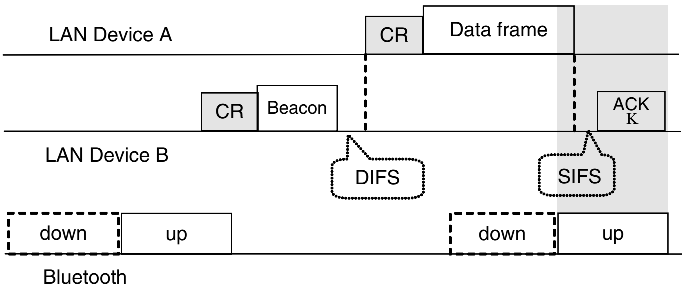

Fig. 8.1 Bluetooth interference timing relative to wireless LAN protocol timing

Figure 8.1 shows, roughly to scale, an on-channel Bluetooth transmission relative to two wireless LAN protocol events: (a) a beacon from one device, and (b) a long frame from another device. The solid outline blocks represent the (uplink) transmissions of the Bluetooth device; the dashed outline blocks represent the down transmissions of the other Bluetooth device, e.g. the Master. If the second Bluetooth transmission frequency falls in the wireless LAN bandwidth, the Bluetooth uplink transmission would destroy the latter part of an incoming data frame or, in case the wireless LAN device sent the data frame, it would destroy the incoming Acknowledgement. Thus, in case of frequency overlap, the "damage window" is the sum of the durations of the wireless LAN unicast cycle and half the Bluetooth dwell time.

Depending on the offered load, the packet loss percentage can be as high as 35%. In practice, wireless LANs are quite able to deal with this kind of interference: throughput is reduced, but the data transport service remains operational throughout. Re-transmissions to recover from frame loss will eat further into the remaining capacity. Retransmission is likely to succeed on the first try, but, for some frames, two or more will be needed. If the average re-transmission rate is 1.5, 5.5 frames have to be sent for every 3 received frames. The resulting throughput loss for different data rates is shown in Fig. 8.2.

Although the throughput loss is dramatic at high loads, it may not be very noticeable at low loads, even if the wireless LAN is used for a streaming-like service such as voice or low rate video. The pain will be mostly felt with high rate video and large file transfers.

The above results depend for a large part on the Bluetooth hit rate, which, in turn, is dependent on the bandwidth of the wireless LAN. Clearly, when 40 MHz channels are used, the loss rate is effectively doubled.

8.1.3.4 Bluetooth as Victim

The preceding analysis shows a large impact of interference caused by nearby a Bluetooth link on a wireless LAN client device, because of frequent "hits" of the wireless LAN channel by a Bluetooth transmission. It stands to reason that the same "hit rate" pattern will apply to wireless LAN interference with the operation of Bluetooth receivers.

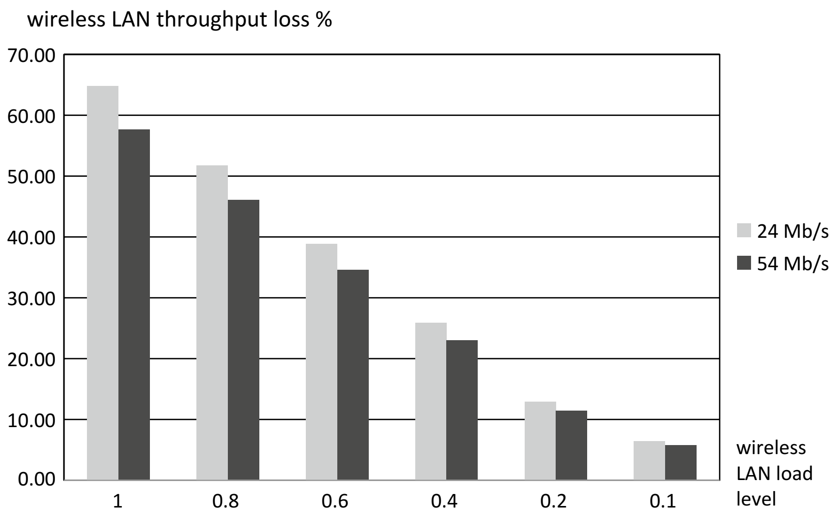

Fig. 8.2 Wireless LAN throughput loss as function of network load

In the following section, we look at three examples: Bluetooth SCO service, Bluetooth ACL service and Bluetooth link set-up, in the context of either a single wireless LAN or multiple wireless LANs interfering with a Bluetooth piconet. The former case models the home network case; the latter, the high density business case.

8.1.3.5 Impact on Bluetooth SCO Service

The Bluetooth SCO service is a connection-oriented service with no provisions for re-transmission. Therefore, every Bluetooth transmission affected by interference means a loss of data, which, in the case of a SCO service, results in very much noticeable loss of fidelity of the voice or video affected.

If access point transmissions interfere with a Bluetooth receiver, the expected hit rate of the wireless LAN transmissions is 400 times per second in the case of a simple home network. In case of a high density network office, this may increase to 1,200 times per second. Note that every time an wireless LAN device receives a frame correctly, it responds with an Ack – and that transmission will interfere with the local Bluetooth receiver. At either of these hit-rates, the impact on the (Bluetooth) user is likely to be dramatic – the service quality drops below the acceptable level.

8.1.3.6 Impact on Bluetooth ACL Service

The Bluetooth ACL service is a connectionless service with some provisions for re-transmission. A Bluetooth transmission affected by interference is retried in case the transmitter receives a NACK from the intended receiver. Given that the NACK

is transmitted on another (hop) frequency, it is not likely that it will be affected. Therefore, re-transmission will recover some of the frame loss but not all of it: if NACK messages get lost, the sender will not be aware of the loss and will not re-transmit. Recovery, if provided at all, has to be performed at a higher level in the stack and that is much less efficient. Therefore, it is safe to assume that also the ACL service will suffer significantly from interference caused by wireless LAN devices. As noted above, wireless LAN client traffic load varies a lot and if only one nearby wireless LAN client causes interference, the overall Bluetooth "data" performance may not be seen as a problem.

8.1.3.7 Bluetooth Channel Adaptation

The regulations for the 2.4 GHz band allow Bluetooth systems to avoid colliding with a nearby wireless LAN through the use of adaptive hopping. However, this modified hop sequence carries a price for the interference reduction: the frequency space available for orthogonal (= independent and therefore non-colliding) hop sets is reduced and, therefore, the number of independently operating piconets is reduced as well. In many cases – e.g. a few Bluetooth devices on a desktop – this will not be a problem. See Sect. 6.6.4 for more details on adaptive hopping.

8.1.3.8 Other Interference Management Measures

The preceding sections sketch a picture of the destructive interference that can occur if two different RF systems share spectrum in the dimensions of time, frequency, and space. Each system incorporates features that facilitate sharing of spectrum with its own kind, but not with other types of system. Since the air interface protocols (TDMA and CSMA) are inherently incompatible, there are two possible solutions in the time domain: changing the protocols or separating their execution in time such that interference does not occur. Otherwise, only separation in frequency or space will reduce or avoid interference.

8.1.3.9 Protocol Level Coordination

Another approach is to coordinate spectrum access at the protocol level: controls can be implemented that make sure that the active receive process in one device is not interfered with by a transmission of the other device. The referenced IEEE co-existence document (see footnote 1 on page 147) does mention such a solution.

In the case of co-located devices, the coordination has to protect both devices during the time that overlap in frequency occurs. Assuming Adaptive Hopping, this overlap covers the skirts of the wireless LAN transmitter and receiver masks plus a margin for the steep fall off of the Bluetooth masks. If the overlapping frequencies are denoted by f_{int} and the range of hop frequencies is f_{hop}, the probability of interference occurring is roughly f_{int}/f_{hop}. Note that f_{int} can be up to 30 MHz. With 20 out of 79

hops closed off by the Adaptive Hopping algorithm, only 59 hops would be left as f_{hop}. Therefore, interference between other wireless devices and Bluetooth devices will occur on every other hop – on average. In practice, that means coordination is required on every hop – that is 1,600 times per second, while a Bluetooth SCO or ACL service is in active use.

8.1.3.10 Application Level Coordination

As the preceding demonstrates, protocol level coordination is not able to provide a satisfactory level of performance of the wireless LAN system – except when its link load is very low or the Bluetooth link load is very low. This leads to the consideration that coordination at the application level may be more effective. Application level coordination involves avoiding co-incident use of both devices by "demanding" applications. Thus, during a Bluetooth voice connection or file transfer, the wireless LAN device is disabled, but it can remain active while Bluetooth is used only for low rate applications. In effect, concurrent sessions are to be avoided if service degradation is expected to be an issue. There are no hard rules for application level coordination since the degree of tolerable service degradation affects the need for session separation.

8.1.4 Summary

Even a cursory examination of the issues related to the sharing of a frequency band by IEEE 802.11 wireless LAN devices and IEEE 802.15 Bluetooth devices shows that there are many reasons why a truly satisfactory way of such spectrum sharing is impossible to achieve. Only when the devices are used in alternating fashion or when either of the two devices is used at a very low rate, users are likely to be unaware of the interference. Since such a situation may well obtain most of the time for most users and therefore implementing both in the same product, be it handheld, portable or fixed, may make good business sense – if only because of the flexibility in connectivity offered by having multiple connection options.

A broader lesson being confirmed by this examination is that sharing spectrum with devices that have not been designed for such sharing is at best inefficient and typically ineffective.

8.2 Wireless LANs and Zigbee

Zigbee has become the commercial label for products based on the IEEE 802.15.4 series of standards, which cover low-rate wireless personal area networks. The *personal* in this term refers to the low power/short range aspect of these specifications rather

than to a particular type of use. Thanks to the efforts of the Zigbee Alliance, Zigbee has become a major contender for in-home energy management applications. Energy management is usually considered to be part of the "smart grid"[7] domain. Zigbee products are available with power levels ranging from 0 to 20 dBm EIRP – the former makes sense for in-home applications, whereas the latter are more frequently used in industrial environments. Although the IEEE 802.15.4 standards cover multiple frequency ranges, in the following only wireless LAN-Zigbee sharing in the 2.4 GHz band is considered.

The Zigbee Physical layer and the Medium Access Control layer have many aspects in common with their counterparts in the IEEE 802.11 specifications. However, the differences determine the overall picture. The channel width of Zigbee is 5 MHz and either BPSK or Offset QPSK modulation is used. The maximum data rate is 250 kbps. Since the typical application space for Zigbee is monitoring and control rather than communications, operation is highly intermittent, the effective duty cycle can vary a factor 100, e.g. between 1% and 0.01%.

8.2.1 *Link Budget Analysis*

8.2.1.1 The Impact of the Channel Asymmetry

The Zigbee standard defines 16 channels of 5 MHz wide. The overlap with the commonly used wireless LAN channel spacing is as follows: Zigbee channels 1 through 4 overlap with wireless LAN channel 1, Zigbee channels 6 through 9 overlap with wireless LAN channel 6 and Zigbee channels 11 through 14 overlap with wireless LAN channel 11. This leaves Zigbee channels 5, 10, 15 and 16 relatively free of wireless LAN interference and vice versa – even at close range (Fig. 8.3).

This suggests that 3 wireless LAN networks and 4 Zigbee networks operating at the same power level can co-exist in the same physical location.

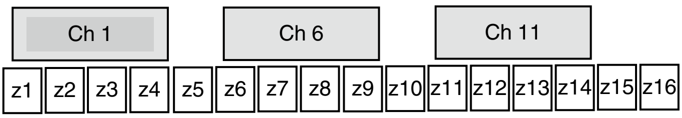

Fig. 8.3 Zigbee and wireless LAN channel overlap

[7]The "smart grid" has been conceived to solve large-scale energy distribution problems.

8.2.1.2 The Impact of the Power Asymmetry

Zigbee devices used for in-home applications and operating at the nominal power output of 1 mW achieve a range of roughly 11 m at 250 kbps.[8]

Table 8.1 gives the RF parameters for a low-power Zigbee device and a "typical" wireless LAN device – the same as used in the wireless LAN Bluetooth sharing analysis.

Wireless LANs normally use *carrier detect* rather than *energy detect* and, because of the large (e.g. 20 dB) difference in thresholds, they ignore other types of devices unless these are operating nearby. This equates to a difference in defer range of a factor 4–8, depending on the type of environment (Table 8.2).

The distances calculated using the above parameters show that the difference in data rate roughly compensates for the difference in transmitter power, Zigbee defers for wireless LANs at a distance (37 m), but that does not protect it from wireless LAN interference (up to 119 m). The wireless LAN defer distance for Zigbee devices is very small compared to the wireless LAN-Zigbee interference distance and, therefore, Zigbee devices are not protected at all by the wireless LAN's defer mechanism. A first observation to be made is that if the wireless LAN drops its data rate to 24 Mb/s, it is largely immune to Zigbee, whereas Zigbee has little margin to make itself immune to wireless LAN interference. Another observation

Table 8.1 Wireless LAN and Zigbee main RF parameters

	Zigbee	Wireless LAN
Transmitter parameters		
Tx power (dBm)	0	17
Antenna gain (dB)	0	0
Bandwidth (MHz)	5	16.5
Receiver parameters		
Bandwidth (MHz)	5	16.5
Antenna gain (dB)	0	0
MUS	–85	–68
Received signal margin	10	22
Energy defer threshold	–75	–65

Table 8.2 Wireless LAN-Zigbee defer and interference distances

Zigbee operating distance	Zigbee–wireless LAN defer distance	Zigbee–wireless LAN interference distance	Wireless LAN operating distance	Wireless LAN–Zigbee defer distance	Wireless LAN–Zigbee interference distance
26	29	37	26	9	119

[8]Assuming a pathloss model of free space up to 4 m, a wall of 10 dB attenuation and 10 dB/octave attenuation beyond the breakpoint.

is that if Zigbee operates at the same power level as the wireless LANs, the differences largely disappear and fair sharing – seen as equal opportunity to use the available spectrum – becomes possible.

Zigbee devices can avoid wireless LAN interference by operating on channels that are not likely to be used by wireless LANs: Zigbee channels 5, 10, 15 and 16. The implication is that Zigbee devices should not actively search for other Zigbee devices on all other channels. However, if and when choosing a free channel is not possible, co-channel operation will be necessary. The above analysis leads to the conclusion that, in case of co-channel operation, any wireless LAN transmission originating anywhere in a fairly large area will interfere with Zigbee operations, whereas the reverse is not true. This result is used in the dynamic analysis below.

8.2.2 Protocol Interaction Analysis

The intent behind the IEEE 802.15.4/Zigbee specification is to provide a robust system for monitoring and control purposes. Distances are short and data rates low – compared to multi-media heavy PC based communications. Thus, 250 kbps was considered adequate and there has been no effort to crank this up to the throughput level possible with wireless LANs. Because robustness was desirable, CSMA/CA was chosen as basic co-existence mechanism rather than frequency hopping employed by Bluetooth. Given a low level of activity, CSMA/CA delivers short response times, as well as graceful degradation of throughput if the medium does get busy.

Monitoring and control applications vary widely in their data demands, so a guess has to be made with regard to a reasonable average. A recent NIST study[9] uses a figure of about 1 b/s for the uplink and 1/8th of that for the downlink. Allowing for a 100-fold increase over this base line figure suggests a peak load of 1,000 bits/s uplink with an average near 1/10th of that. Using the average value and assuming a frame overhead of at least 42 bytes gives one 1,440 bit message every second. At the basic rate of 250 kbps of Zigbee, this takes 5.76 ms. A downlink message of 125 bits with 42 bytes of overhead would occupy roughly 1/3rd of that or ~2 ms. The Zigbee link capacity is therefore near 125 message pairs per second. Assuming a 10-unit network the actual load works out at 12.5 messages per device. This seems rather much for simple energy management purposes.

As shown above (see Table 7.6), the effective throughput of an isolated wireless LAN running at 48 Mb/s in short pre-amble mode and average frame size is 46% = 19.29 Mb/s. Watching a video in standard resolution requires about 3 Mb/s which is 13.6% of the link's net capacity. Adding allowance for other traffic, the peak load of a wireless LAN in a home network can be put at 25%. Packet duration will be in the range of 125 μs and frame interval in the range of 500 μs. Since a Zigbee frame

[9] See e.g. the NIST website. A good introduction is available as collaborate.nist.gov/twiki-sggrid/… /**Smart**Grid/…/LTE_**Smart**Grid_Analysis.ppt.

exchange takes about 7 ms, the wireless LAN traffic will disrupt co-channel Zigbee traffic with high probability: during the Zigbee frame exchange, there may be as much as 14 wireless LAN transmissions – in this scenario. In other words, there is no way that a protocol adaptation of Zigbee can change this.

8.2.3 Summary

The preceding cursory analysis considers only the basic facts about wireless LAN and Zigbee RF properties and protocol behavior. Nonetheless, it is clear that the large asymmetries in power and defer distances make co-channel operation of wireless LANs and Zigbee within interference range impossible. Only separation in the frequency domain can address this inherent incompatibility. From this point of view, it is fortunate that the IEEE 802.15.4 Working Group chose a channel scheme that provides at least four free channels for Zigbee operations, even in an environment with heavy wireless LAN usage.

8.3 Wireless LANs and Ultra-Wide-Band Technology

Even though Ultra-wide-Band technology has never seen the rate of adoption its early proponents foresaw, it is worth looking into the spectrum sharing between wireless LANs and Ultra-Wide-Band systems. The latter present an extreme approach to spectrum sharing: very low power at very large bandwidths. For lower frequencies, there is no dynamic sharing mechanism involved, but at frequencies above 3.1 GHz Ultra-Wide-Band devices have to implement a "detect and avoid" (DAA) mechanism to protect radar and broadband wireless access systems.[10] This mechanism is a bit like the DFS mechanism used by wireless LANs to protect radars in the 5 GHz band.

Ultra-Wide-Band technology is used in many applications; there are chips and other components on the market that find application for various purposes and in different environments – including industrial communications systems that leverage its interference resistance. Ultra-Wide-Band has been adopted by the USB community.[11]

Ultra-Wide-Band technology is also in use as ground/wall penetrating radar; that type application is not considered here.

[10] See ECC Report 120: Technical Requirements for UWB DAA (detect and avoid) devices to ensure the protection of radiolocation services in the bands 3.1–3.4 GHz and 8.5–9 GHz and BWA terminals in the band 3.4–4.2 GHz.

[11] See www.USB.org. It states that the Wireless USB performance is targeted at 480 Mbps at 3 m and 110 Mbps at 10 m.

8.3.1 Wireless LAN Interference and UWB Emission Limits

The FCC identified 3,100–10,600 MHz as the Ultra-Wide-Band "transmission band." Within that band, a peak power limit of 0 dBm/50 MHz applies. In addition, Table 8.3 applies.

Here we address the frequency bands used by wireless LANs: the 1,990–3,100 MHz range and the 3,100–10,600 MHz range.

It should be noted that the −41.3 dBm level is the same as the spurious emission limit for unintentional radiators, such as PCs and vacuum cleaners. In effect, this means that the peak level of unwanted emissions is made into a generic, potentially full spectrum limit of an Ultra-Wide-Band transmitter. To put this into perspective: a conventional wide band transmitter may have one or a few spurious emission spikes in its spectrum and that is considered acceptable since the probability of overlap with a victim receiver is not great. However, an Ultra-Wide-Band device is allowed to fill that whole frequency range with a much higher level of total energy: −41.3 dBm/MHz = −14.3 dBm/500 MHz. In order to assure that pulsed Ultra-Wide-Band systems would not cause very high peaks of interference, the FCC added a peak power limit of 0 dBm/50 MHz, which the FCC equates with a 16.8 dB allowance for pulsed systems.

In the case of wireless LAN victim receivers, the bandwidth ratio clearly plays a role. For a 2.4 GHz wireless LAN, the threat comes from devices that can output up to −51.3 dBm/MHz. For indoor scenarios, this amounts to −39.1 and −35.5 dBm in-band power. For the 5.4 GHz band, these levels are about 10 dB higher (see Table 8.4).

However, taking into account the 16.8 dB/MHz limit changes the picture significantly (Table 8.5).

Since the signal attenuation over short distances tends towards 6 dB per octave, the interference power at 2 m distances is easily calculated (Tables 8.6 and 8.7).

Since signal attenuation at 5 GHz is 6 dB more than at 2.4 GHz, the degradation of the SNIR caused by Ultra-Wide-Band devices at 5 GHz is about 2 dB worse than at 2.4 GHz. Ignoring that difference and looking at the effect of average power, the interference level for an wireless LAN receiver is about −90 dBm for a system using 20 MHz channels. The noise floor of such a receiver is −102 dBm; taking into account

Table 8.3 FCC: Ultra-Wide-Band average transmission power levels

Frequency (MHz)	EIRP (dBm/MHz), indoor	Idem, handheld
960–1,610	−75.3	−75.3
1,164–1,240	−85.3	−85.3
1,559–1,610	−85.3	−85.3
1,610–1,990	−53.3	−63.3
1,990–3,100	−51.3	−61.3
3,100–10,600	−41.3	−41.3
Above 10,600	−51.3	−61.3

Table 8.4 Wireless LAN Interference burden based on Ultra-Wide-Band average power

Wireless LAN bandwidth	Frequency (MHz) 1,990–3,100	3,100–10,600
2.4 GHz, 16.5 MHz channel	−39.1 dBm	
2.4 GHz, 38 MHz	−35.5 dBm	
5.4 GHz, 16.5 MHz		−29.1 dBm
5.4 GHz, 38 MHz		−25.4 dBm
5.4 GHz, 76 MHz		−22.4 dBm

Table 8.5 Wireless LAN Interference burden based on Ultra-Wide-Band peak power

Wireless LAN bandwidth	Frequency (MHz) 1,990–3,100	3,100–10,600
2.4 GHz, 16.5 MHz	−39.1 dBm	
2.4 GHz, 38 MHz	−35.5 dBm	
5.4 GHz, 16.5 MHz		−11.7 dBm
5.4 GHz, 38 MHz		− 9.4 dBm
5.4 GHz, 76 MHz		− 7.8 dBm

Table 8.6 Wireless LAN interference at 2 m distance based on Ultra-Wide-Band average power

Wireless LAN bandwidth	Frequency (MHz) 1,990–3,100	3,100–10,600
2.4 GHz, 16.5 MHz	−91.9 dBm	
2.4 GHz, 38 MHz	−89.9 dBm	
5.4 GHz, 16.5 MHz		−87.9 dBm
5.4 GHz, 38 MHz		−84.9 dBm
5.4 GHz, 76 MHz		−81.9 dBm

Table 8.7 Wireless LAN Interference at 2 m based on Ultra-Wide-Band peak power

Wireless LAN bandwidth	Frequency (MHz) 1,990–3,100	3,100–10,600
5.4 GHz, 16.5 MHz		−71.1 dBm
5.4 GHz, 38 MHz		−69.1 dBm
5.4 GHz, 76 MHz		−67.5 dBm

a generous implementation margin of 6 dB, the net impairment caused by a single Ultra-Wide-Band device at 2 m distance is at least 6 dB. This halves the maximum operating distance of the wireless LAN for any data rate. Conversely, the capacity loss is roughly a factor 2: halving the bit rate improves the SNIR by 6 dB. This result generalizes to all data rates, including MIMO rates. The peak power values allowed are 16.8 dB higher and the impact on wireless LAN operations is correspondingly

Table 8.8 ECC limits for Ultra-Wide-Band transmitters[a,b]

Frequency range	Maximum mean EIRP spectral density	Maximum peak EIRP (measured in 50 MHz)	FCC – equivalent for handheld devices
Below 1.6 GHz	–90 dBm/MHz	–50 dBm	–59.1 dBm
1.6–2.7 GHz	–85 dBm/MHz	–45 dBm	–36.5 dBm
2.7–3.4 GHz	–70 dBm/MHz	–36 dBm	–34.5 dBm
3.4–3.8 GHz	–80 dBm/MHz	–40 dBm	–26.5 dBm
3.8–4.2 GHz	–70 dBm/MHz	–30 dBm	
4.2–4.8 GHz	–70 dBm/MHz	–30 dBm	
4.8–6 GHz	–70 dBm/MHz	–30 dBm	
6–9.1 GHz	–41.3 dBm/MHz	0 dBm	
9.1–10.6 GHz	–65 dBm/MHz	–25 dBm	
Above 10.6 GHz	–85 dBm/MHz	–45 dBm	–36.5 dBm

[a]UWB devices placed on the market before 31 December 2010 are permitted to operate in the frequency band 4.2–4.8 GHz with a maximum mean EIRP spectral density of –41.3 dBm/MHz and a maximum peak e.i.r.p of 0 dBm measured in 50 MHz

[b]In case of devices installed in road and rail vehicles, operation at 6 to 8.5 GHz is subject to the implementation of Transmit Power Control (TPC) with a range of 12 dB with respect to the maximum permitted radiated power. If no TPC is implemented, the maximum mean EIRP spectral density is –53.3 dBm/MHz

higher. However, the peak power point of the Ultra-Wide-Band device many not overlap the wireless LAN channel and the pulsed nature of the Ultra-Wide-Band signal will reduce the impact of the interference. However, adding another 6 dB to the effective interference level is warranted. The resulting range or capacity loss is another factor two. For wireless LANs using MIMO type transmission, the effects will not be different. These results are not unexpected and show that those who opposed the Ultra-Wide-Band rulemaking of the FCC had a point.

Economically, the case is less clear cut: the capacity loss of an indoor wireless LAN may not be noticeable in all cases and, if it is, the cause is an Ultra-Wide-Band device under control of the wireless LAN owner – it is his/her problem, literally. In the outdoor case, that control does not apply: the mobile device's loss of throughput may be caused by an Ultra-Wide-Band of a stranger and the victim would have no recourse. Therefore, Ultra-Wide-Band devices are subject to power restrictions.

The EU went through an extensive debate on the merits and demerits of Ultra-Wide-Band technology – much like the US did. The resulting rulemaking, although it reflects EU particulars, is roughly comparable (see Table 8.8).

In summary, this section shows that Ultra-Wide-Band transmitters at close range affect wireless LAN devices – and by extension many other commodity wireless devices – in the sense that they either reduce the operating range or the achievable data rate. The latter point is important for high data rate applications which require a high SNIR. However, the market share of Ultra-Wide-Band communications technology has remained small and applications tend to address specific environments rather than broadly used consumer gear.

Chapter 9
Sharing with Primary Spectrum Users

RF spectrum has to be shared, not only between systems and applications of similar purpose and regulatory status, but also between primary spectrum users and commodity wireless systems. The former are deemed to merit protection from interference caused by the latter. This is an example of vertical spectrum sharing between dissimilar systems. This chapter provides three examples: wireless LANs and C-band radar systems, wireless LANs and 4G/LTE systems, and license exempt systems operating in the TV White Space spectrum. The chapter closes with considerations on the use of Cognitive Radio techniques for vertical spectrum sharing.

9.1 The 5 GHz Band – Radar Systems and Wireless LANs

This section provides an analysis of the vertical spectrum sharing regime in the 5 GHz band that allows wireless LAN and other equipment to operate in this band, provided they do not cause interference to radar systems. The following sections provide the general picture; Appendix B provides more details and also gives the regulatory requirements and compliance criteria.

9.1.1 Background

In the mid 1990s, the regulators and the nascent wireless LAN industry converged on a set of rules for the 2.4 GHz ISM band that allowed wireless LAN technology to take off and broaden into advanced modulation schemes, such as OFDM, that went well beyond what was possible under the spread spectrum rules originally conceived for this band.

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_9,

With the increased throughput capabilities came the possibility of broadening the applications to include not just wireless access to the internet, but also consumer-oriented applications like wireless video links for use in the home. The prospect of increased density of use and large SNR margins required for high rate links pointed towards the need for more spectrum than was available in the 2.4 GHz band. The quest for more spectrum focused early on at the 5 GHz band. It is partially allocated to the Mobile Satellite Service (uplink), but otherwise the domain of radars and related equipment – notably the Microwave Landing System (MLS), ground-based radars, airborne radars and satellite-based passive and active sensors – collectively known as the Earth Exploration Satellite Service (EESS).

Radar systems have two properties that make them sensitive to interfering signals: very high antenna gain and very sensitive receivers. A typical radar receiver may be 1 MHz wide and its sensitivity is only limited by the noise figure of the implementation. A 4 dB noise figure puts the effective sensitivity at −110 dBm. Add a 45 dB antenna and the effective sensitivity drops to −155 dBm. With such a level of sensitivity, radars will pick up wireless LAN interference at long distances. Figure 9.1[1] shows the impact of an outdoor wireless LAN operating at 18 dBm radiated power at a distance of 6.4 km.

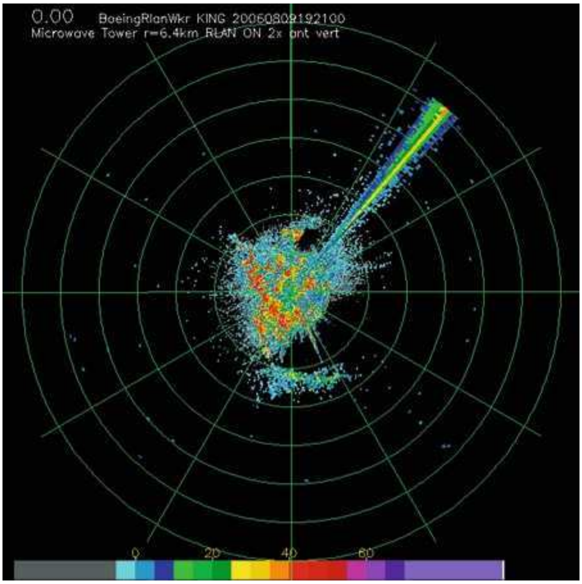

Fig. 9.1 Example of wireless LAN interference into a weather radar

[1]From Boeing document D6-83753 – [DFS and airborne wireless LANs in the 5 GHz range]. Reproduced with the kind permission of the Boeing Company and Environment Canada.

Note: The high antenna gain is also a benefit in the context of radar detection by wireless LAN: the radar energy is concentrated in space and time and this facilitates reliable detection. The detectability of radar signals and the flexibility of wireless LANs together made it possible to devise the first dynamic vertical spectrum sharing scheme: Dynamic Frequency Selection or DFS. The concept of DFS is simple: if a wireless LAN detects a radar signal, it has to move its operating channel to another frequency so that its transmissions will not interfere with the radar's operation. That simple concept was the basis for the DFS requirements and the test patterns developed in ITU-R Recommendation M.1652 and its descendants. The practice of DFS is far from simple: radar signals vary in the spatial, the frequency and the time domain and detection is affected by the wireless LAN's own operations. Conversely, the radar's behavior, the absolute distance, and the pathloss between wireless LAN and a radar affect the interference seen by that radar.

9.1.2 Radar Properties and DFS Performance

In order to understand how DFS has to work and how it can be tested, one has to know the key parameters of the radar systems to be protected: power output, antenna gain, rotation/scan rates, pulse modes and rates, and the I/N protection margin. The following summarizes these and other factors that affect radar detection and wireless LAN, causing interference. Appendix B.1 treats the same material in more detail.

9.1.2.1 Radar Types and Radiation Patterns

Many radars which operate in the 5 GHz band are used for weather analysis, navigation and/or military target acquisition purposes. There are fixed and mobile ground-based radars of this type, as well as shipborne radars. These are all high power systems with power outputs in the range of 50–150 kW. Antenna gain is in the range of 40–50 dB and rotation speed is typically low, e.g. 6 to 12 revolutions per minute. Pulse duration is in the range of .5–5 μs and repetition rates are commensurate with the long range – 250 pps for older radars. Newer designs use faster, but multiple pulse rates which avoids the range ambiguity associated with a single pulse rate. Current meteorological radars make use of pulse widths between 0.5 and 2.5 μs. The typical PRF used ranges from 250 to 1,200 pps and is frequently staggered.[2] In addition to these large fixed radars, there are some mobile tactical radars that operate in the 5 GHz band. They are typically used for local air defense and battlefield surveillance radars. Due to size restrictions, antenna gain is usually lower, e.g. 30–36 dB; they use lower power transmitters, higher pulse rates and high antenna rotation speeds. Advanced digital signal processing is employed to counter interference and jamming.

[2] See page 203 under "Pulse detection statistics".

9.1.2.2 The Detection Threshold

The detection threshold for DFS was determined by simulations[3] performed by experts of the NTIA during the preparations for WRC 2003. These simulations included a large number of parameters that describe the density distribution and location of wireless LANs in a hypothetical city on the coast. This allowed maritime radars to be accounted for as well. The resulting figures of −62 dBm for indoor wireless LANs and −64 dBm for outdoor wireless LANs became the widely agreed basis for the radar-wireless LAN co-existence regime.

9.1.2.3 Antenna Gain and Apparent Radar Beam Shape

Antenna gain and directionality are key elements of radar systems, regardless of the type of antenna used – physical or electronic.

Nominal radar beam shape is given in terms of the half power beamwidth for the main beam and major sidelobes. This nominal figure is useful to determine the spatial resolution of a radar at extreme range or minimal target size, but it says very little about the signature of a radar signal as seen by a DFS detector that may be located rather close to a radar.

In general, the link budget for the detection of a radar by a DFS detector is much greater than the link budget that determines wireless LAN interference seen by the radar. This is shown in Fig. 9.2 below for a high power radar system of 250 kW output power and a wireless LAN of 200 mW radiated output.

This link budget difference is determined by the EIRP of the transmitters and the thresholds of the receivers and not by parameters of antenna gain, pathloss, etc.; these affect the actual distances involved. The radar horizon determines whether a

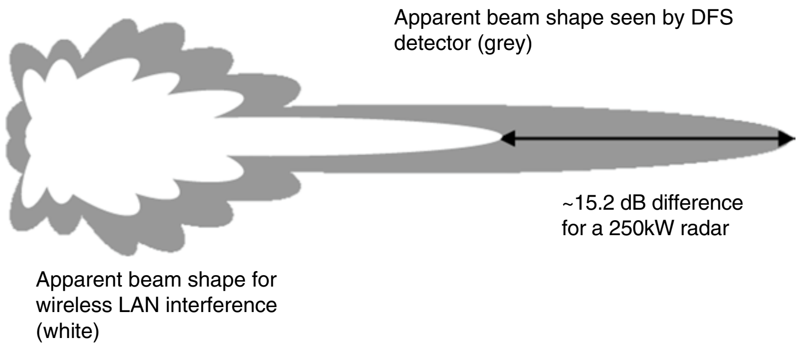

Fig. 9.2 Apparent beam shape differences

[3] See Report 09–461 available from http://www.ntia.doc.gov/osmhome/reports.html.

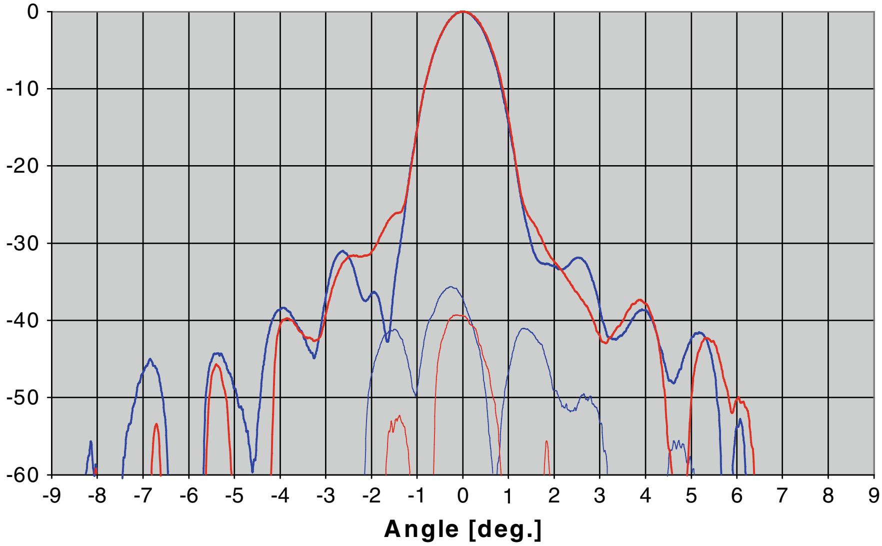

Fig. 9.3 Beam shape of a modern weather radar (dBr/deg) (Reproduced with kind permission of the Vaisala company)

radar is visible to a DFS detector. Within this "footprint" of the radar, the signal strength is very high relative to the regulatory detection threshold: it can be as much as 40 dB higher. This means that the DFS detector sees the 40 dB beamwidth, which is much wider than the nominal beamwidth.

As Fig. 9.3 above shows, the 40 dB beamwidth is nearly 6° wide – that is about 7 times as much as the nominal beamwidth. This makes the radar signal easily detectable.

9.1.2.4 Observed Burst Length

The rotation and scanning motions of radar antennas limit the time a given point in space is illuminated by the radar beam. The observed burst length is given by the pulse repetition rate (PPR) divided by the dwell time of the beam: PPS*°BW/(6*RPM), in which BW is the effective beamwidth.

An analysis covering nearly all weather radars in Europe[4] showed that, except for a few old radars only used for storm warning, the number of pulses seen by the DFS detector would be in the range of 43–145 pulses per "radar sweep." This is adequate to assure a very high detection probability.

[4] See ECC Report 140, source ERO.

9.1.2.5 Pulse Detection Statistics

In practice, the DFS detection efficiency is determined by the statistics of the pulse patterns (intervals and burst length) and by operational conditions such as the wireless LAN busy level. Radars use different pulse patterns depending on their purpose; relevant from a detection point of view is that some use pulse patterns with two or three pulse intervals. These intervals may be applied to individual pulses, as well as to bursts of pulses. This is known respectively as pulse staggering and burst staggering.

Figure 9.4 shows how detection probability varies with the number of pulses per burst and with the detection probability for individual pulses – in this case, for a comfortable false alarm threshold of 4 pulses.

Staggered pulse patterns affect the detection probability in different ways: burst staggering results in bursts of pulses at some constant spacing and, therefore, detection probability is the same as for constant pulse rates. Pulse staggering requires more pulses because to the reduced redundancy in the pulse train: a 6 pulse false alarm threshold (= 50% more) provides additional redundancy to compensate for this. Given the high pulse rates of staggered PRF radars, this 50% is easily reached. In other words, whereas an analysis based on nominal radar data suggests difficulties in detecting certain radars because of the short nominal burst lengths seen by a DFS detector, the actual burst lengths observed from real radars show ample margin and, therefore, high detection probabilities.

The preceding also indicates that radar operators have considerable influence on the detectability of their systems: higher PRFs mean increased detectability.

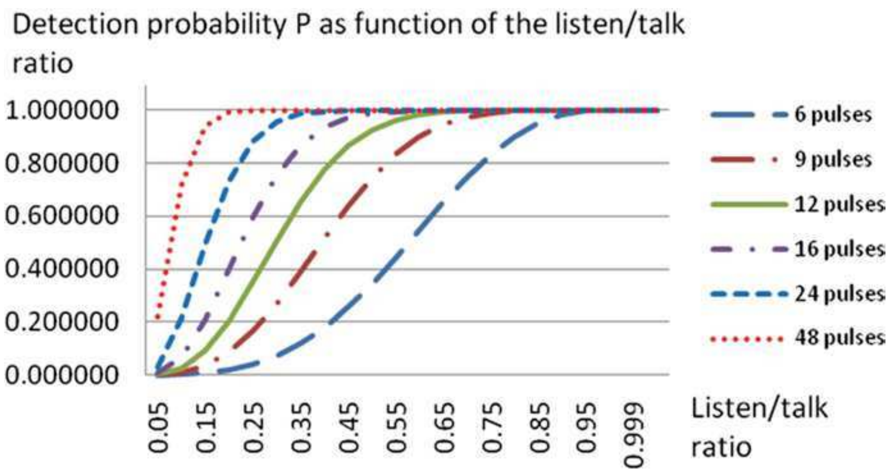

Fig. 9.4 Radar Detection probability under different conditions

9.1.2.6 Pulse Shape and Spectrum

Radar pulses tend to be short: from .5 to 5 μs and rising and falling edges tend to be commensurate: 1–10%. Steep slopes give rise to a wide frequency spectrum of short duration. Short rise times, necessary for certain radar modes of operation, cause a wider spectrum and, therefore, cause radar detection over a wider spectrum than the wireless LAN bandwidth. The implication is that wireless LANs will detect radars at frequencies at which they may not cause interference. Increased wireless LAN bandwidth – like used in MIMO systems – aggravates this difference – to the disadvantage of the wireless LANs.

9.1.2.7 Radar Unwanted Emissions

Whereas the above pulse spectrum is considered part of the necessary bandwidth of a radar – necessary in terms of the radar's mission – the unwanted emissions that extend further out from the fundamental are not necessary, but they are consequence of design and implementation choices.

Radar spectra vary with the design of the transmitter amplifier – klystrons tend to be cleaner than magnetrons and filters can be applied to reduce the "skirts" without affecting the necessary bandwidth much at all. Because of the wide bandwidth of the wireless LAN receiver, the apparent strength of the spurious emissions is well above the DFS threshold, even at considerable distances – more than enough to trigger the DFS channel blocking process.

However, none of the detections caused by outlying spurious peaks would correspond to real interference threats, because the radar may not be sensitive at those frequencies. The implication is that a DFS detector could see a number of "occupied channels" where as only one occupied channel might be adequate to protect the radar.

9.1.2.8 Radar Scan Patterns

Radar scan patterns vary with the application and with advancing technical development. To yield useful information for the radar receiver, the illumination of the to be detected objects has to deliver a certain amount of energy. This amount of energy is determined not only by the transmitter's power output and antenna gain, but also by the scan rate of the antenna, the pulse repetition rate of the transmitter and the pulse width. Given the need to deliver enough energy (measured in pulses and power per pulse at a given operating frequency) to a given point in space, reliable radar detection is generally feasible.

However, little is known about military radars. Many are still based on the basic model of a rotating antenna, combined with a vertical fan beam or vertical scanning. Some combine search and tracking functions. More recent designs combine electronic beam steering and frequency hopping, which improves jamming resistance and reduces detectability. Pulse or burst staggering may be used to improve range discrimination.

Table 9.1 Factors that affect radar detection

Factor	Depends on	Impact
DFS threshold	Regulation	Higher threshold = less false alarms; lower threshold = larger margin between detection and interference link budget; better detection of low power radars
Radar antenna beamwidth	Radar mission and design	Higher directivity = lower pulse count, more difficult to detect
Radar beam dwell time	Rotation rate and beamwidth	Higher dwell time = faster and more reliable detection and vice versa
Pulse rate/pulse patterns	Radar mission and design	Higher pulse rate = more reliable detection and vice versa; staggered pulse pattern reduce detection probability somewhat
Wireless LAN busy level	Local conditions	Higher level reduces detection probability for a given burst length – non-linear effect
Pulse shape	Propagation conditions	Hardly affects detection
Pulse spectrum	Pulse rise time	Steeper rise times = more false (off channel) detections that limit available spectrum within the radar horizon
Radar unwanted emissions	Radar design	More spurious means more false radar detections – limits available spectrum within the radar horizon
Radar scan pattern	Radar mission	May affect detection opportunities and reduces interference opportunities

The most advanced civilian radars are weather radars. These combine staggered PRFs, multiple polarizations and helical scan patterns. The latter has consequences for DFS operations. The example in Appendix B.1 shows that such a scan pattern may visit the horizon only for a few revolutions per scan cycle. DFS detection procedures must take this into account.

9.1.2.9 Summary

The preceding sections show that radar detection is a complex technical issue that involves many factors. Some work in favor of detection, some hinder detection, and some cause false detections Table 9.1 provides a summary.

9.1.3 Radar Sensitivity and Wireless LAN Interference

For civilian radar systems, the essential receiver parameters are known and the digital signal processing applied by weather radars is public material. These parameters largely determine the protection criteria and allow a risk assessment. Since not many of the key parameters of military radars are known, it is difficult to describe how wireless LAN emission affects these radars. Therefore, one has to use the basic

RF parameters in interference analysis and leave the limited knowledge available about digital signal processing – such as electronic countermeasure capabilities – out of consideration.

Because radar systems are deployed all over the RF spectrum and because, in most cases, they share such spectrum with other systems, the ITU-R has developed protection criteria for radar receivers. The basic measure is the I/N ratio: the difference between the receiver's noise floor and the interference signal. The ITU-R applies two rules for setting the I/N margin for spectrum sharing: I/N = −6 dB is used between co-primary services, I/N = −10 dB is used between a primary and a secondary service. In the case of wireless LANs in the 5 GHz band, the two services (radiodetermination and mobile) are co-primary and, therefore, the −6 dB should be used. This was the case also for the assessment of the sharing between MSS and wireless LANs in the lower 5 GHz range. The Radar-wireless LAN co-existence simulations done in the preparations for WRC 2003 also used this value.

9.1.3.1 Radar Sensitivity and Selectivity[5]

Radar sensitivity must be high for all radar applications, but weather radars, notably those used to detect wind shear and direction, probably represent the highest performing civilian radar receivers. The data below are representative of many radars of this type (Table 9.2).

On the basis of the above and the default protection criterion of I/N −6 dB, the interference threshold for this class of radars is −118.6 dBm/MHz. Due to the need for tuning the operating frequency to local conditions, the front-end bandpass filter of these radars is about 10 MHz wide at the −3 dB points and about 40 MHz wide at the −40 dB level.

The filter slope of 27 dB/MHz, together with the delta between the detection link budget and the interference link budget, determines the minimum frequency separation needed between the operating frequencies of a radar and wireless LAN devices within interference range. This suggests that the wireless LAN transmitter should not operate closer to a radar transmitter than 15 MHz + a half RLAN channel width. In practice, this comes down to 30 MHz centre-to-centre frequency separation.

Table 9.2 Typical weather radar receiver data

Parameter	Value
Front end filter bandwidth	10 MHz
I/F receiver bandwidth	1 MHz
Receiver noise figure	1.4 dB
Receiver noise floor	−112.6 dBm

[5] From NTIA Report 09-461.

9.1.3.2 Link Budgets and Separation Distances

As noted above, the ratio between the link budgets for detection and interference determines whether a given wireless LAN will be able to cause interference: $Tx_{radar} + Threshold_{rlan} > Tx_{rlan} + RxInt_{radar}$. If this inequality holds, the wireless LAN will detect the radar and go off to another channel. Pathloss antenna gain affects the distances involved: for high power/high gain systems the interference zone stretches up to the local horizon. For low power systems, the interference zone may be small – e.g. a few kilometers – so that the number of potential wireless LAN interferers is almost negligible given the operational use of such radars – e.g. on the battlefield.

The interference link budget is determined by the operating frequency, the wireless LAN's EIRP and the radar/wireless LAN bandwidth ratio, the radar antenna gain and the radar's protection level. For an indoor wireless LAN operating at 200 mW EIRP in the 5 GHz range, this link budget is given by the sum of:

- the wireless LAN EIRP
- the antenna gain, the wall loss
- the radar/wireless LAN bandwidth ratio
- the net antenna gain of the radar antenna
- the interference threshold of the radar
- the frequency dependent 1 m loss. In this case: 23 + 0 − 14 − 12.2 + 48 + 118.6 − 47.4 = 116 dB

This corresponds to a separation distance of 60.5 km in free space – which is beyond the local radar horizon. For −40 dB on the slope of the radar's receiver front-end filter, this distance drops to 4.73 km. The detection range in the latter case is determined by the radar's transmission power roll off. Although precise data are lacking, the literature[6] suggests an effective −40 dB bandwidth of some 40 MHz and a −60 dB bandwidth of 60–80 MHz. The former would determine the detectability of the radar in adjacent channel cases. At −40 dBr and a 6 dB detection margin DFS would detect the radar at a distance of 13.7 km, i.e. more than the distance at which it could cause some interference. The latter would determine the detectability of the radar at larger frequency separations: at −60 dBr, the wireless LAN would detect the radar at 2.12 km. The required separation distance is determined by the radar's front-end filter. At 40 MHz its attenuation should exceed the −40 dB by at least 20 dB. This drops the interference distance to a very short 51 m.

9.1.3.3 Interference into Radar Antenna Sidelobes

The above analysis and calculations were all based on the main beam gain (−3 dB beamwidth) of the radar. Since equal antenna gain values apply to the detection path and to the interference path, the analysis holds also for the sidelobes of the radar antenna.

[6] See NTIA TM-05-431, available at http://www.its.bldrdoc.gov/pub/ntia-rpt/05-431/05-431.pdf.

9.1.3.4 Low Power and Wide Band Radars

Low power and wideband radars have in common that their power density is lower than the conventional high power radars discussed in the preceding sections. This changes the detection/interference link budget ratio. As we have seen, this ratio is about 15 dB for a 250 kW radar and a 3 dB DFS detection margin. For a 12.5 kW radar, this ratio drops to 2 dB. For a 6 dB detection margin, it would drop below zero: detection and interference distances remain roughly equal. Since low-power radars typically have smaller – and therefore lower gain – antennas, these distances are shorter as for high gain antennas, but both would still be in the same range as the radar horizon. Interference probability would effectively be nil.

For the off-channel interference discussed in Sect. 3.3.3, this does not apply; assuming that the effective radar transmit filter slope roughly equals the slope of its receiver front-end, the above distances drop to 1.2 and 1.25 km respectively even for the I/N = −6 dB protection margin. At such short distances, there will not be many wireless LANs around to cause any interference.

The above applies also for wideband radars – some apply frequency “chirping” over a bandwidth of 5 or more MHz. This widens the transmitted bandwidth, as well as the receiver front-end filter bandwidth; these effects should roughly cancel each other.

9.1.3.5 Outdoor Wireless LANs

From an interference analysis point of view, the major difference between indoor and outdoor wireless LANs is that the latter are allowed to use 7 dB more radiated power and that their signals are not attenuated by walls. Since both factors affect the interference path as well as the detection path, there is no difference in the protection parameters. However, there is a major difference in the interference level – if interference occurs. Therefore, outdoor wireless LANs present a major case for tight enforcement of the DFS compliance criteria. Tighter technical requirements do not deter malicious intent or compensate for lack of respect for the law.

9.1.3.6 Wireless LANs On-Board Aircraft

The interference potential of wireless LANs on-board aircraft into military radars and weather radars was analyzed extensively in ECC Report 140.[7] During the preparation of this Report, the importance of the apparent radar beamwidth – as opposed to the nominal radar beamwidth – was discovered. This ECC Report showed that, thanks to the very high detection probability for all weather radars, the chance of interference from wireless LANs on board aircraft into these radars would be very remote.

[7] It can be obtained from the ERO website: http://www.erodocdb.dk/doks/.

9.1.3.7 Wireless LANs in Surface Transportation Systems

Although the simulations that underlie the DFS detection threshold choice assumed airborne radars, the current DFS compliance criteria assume a fully stationary wireless LAN as threat model. Given the locations and "footprint size" of major radar systems, the increasing penetration of wireless LANs in transportation systems would suggest the potential for increased risk of wireless LAN interference. This specific consideration was one reason why the Canadian Administration kept the 5,600–5,650 MHz band closed for wireless LANs.

However, analysis shows there is ample margin for a solution that leverages the possibility of instant response to radar detection so as to suppress wireless LAN transmissions in real time.

There are two threat scenarios that are relevant: the multi-lane highway that passes in full view of an elevated weather radar and a weather radar with busy city in its footprint. In the first case, it is the intensity of the interference that could cause problems; in the second case, the potential threat comes from the volume of wireless LANs.

The first case, the highway, is characterized by a relatively short distance between radar and highway. Therefore, the radar's signal seen by DFS detectors is strong and the apparent beamwidth large – as much as 6–7 times the nominal beamwidth. As noted above, this opens up the possibility for lateral detection: the large number of pulses seen by the DFS detector makes it possible to operate the wireless LAN DFS function in a listen-before-talk mode; radar detection would cause suppression of the wireless LANs transmission, while the radar antenna points at the wireless LAN device only. The timing involved is shown in Fig. 9.5 for a weather radar rotating at 6 RPM.

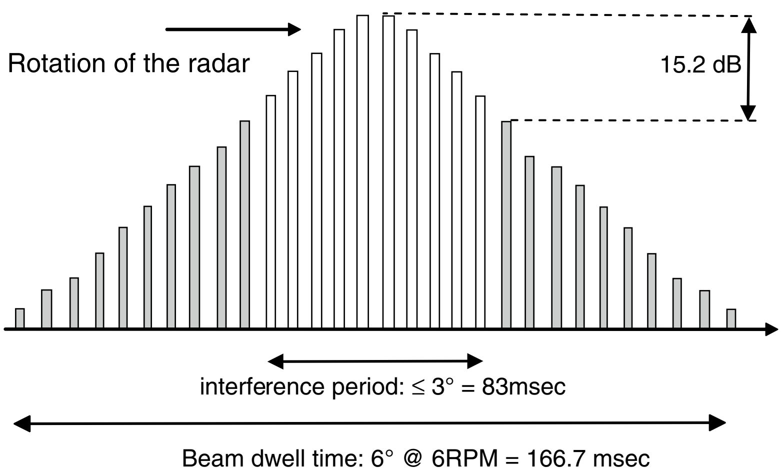

Fig. 9.5 Timing of lateral DFS detection

Table 9.3 Summary of wireless LAN interference factors

Factor	Depends on	Impact
Radar protection criteria	Regulatory rules	Tighter criteria mean more problems meeting them
Radar sensitivity and selectivity	Radar design	More sensitivity is no problem and increased selectivity would help reducing the frequency separation in adjacent channel interference
Pathloss and separation distances	Environment	More pathloss means shorter separation distances
Radar antenna sidelobes	Antenna design	No impact on wireless LAN interference
Low power radars	Radar design	Reduces the sharing margins – needs lower DFS threshold
Outdoor wireless LANs	Regulatory rules	Higher intensity of interference, not higher risk of interference
Wireless LANs on board aircraft	Regulatory rules	None to date
Wireless LANs in surface transportation	Not applicable	None to date

In this case, the number of pulses available for DFS detection in the first quarter of the beam dwell time depends on the radar's PRF. For a fast PRF of 1,200 pps, that number is about 50 pulses. Such radars would be easily detected and even those operating at 400 pps would be well protected. The same argument applies to DFS detectors aboard vehicles elsewhere in view of a radar site. However, the DFS compliance criteria do not cover this mode of detection and response behavior. On the contrary, the FCC's DFS compliance criteria allow for only 260 ms of coordination transmissions following radar detection, and this encourages designers to "immediately" start off the channel change process rather than waiting for the radar beam to pass.

The second case, the large number of wireless LANs in vehicles spread over an urban area, is different. However, much of the large number of vehicles will be in the radar shadow created by buildings and other structures and therefore invisible. Small "pockets of mutual visibility" might exist and here the detection scenario sketched in the preceding paragraph applies. Therefore, the question is if those pockets would be enough to cause interference. Simulation and experiment may be necessary to settle this more complex question.

Table 9.3 summarizes the preceding sections.

9.1.4 DFS 2.0: Improving DFS Compliance Criteria

The current DFS compliance criteria focus on a narrow range of radar behaviors. Because radars with behaviors outside this narrow range might suffer interference, this presents a risk for the incumbents as well as the wireless LAN industry.

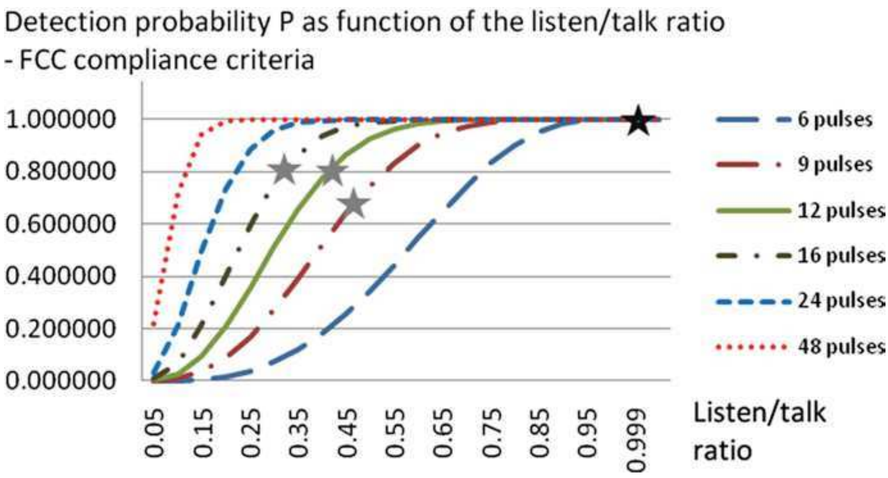

Fig. 9.6 DFS compliance validation points

Thus, the main requirement for new compliance criteria is to increase the coverage of the criteria over a wider range of detection parameters that correspond better to real world conditions. Figure 9.6 shows how detection probability varies with the number of pulses per burst and with the detection probability for individual pulses. The stars roughly mark the "test points" given by the FCC's criteria: the lower gray star marks the position of the frequency hopper test criteria, the other gray stars marks the maximum requirement for the types 1–4 radars. The rightmost star marks the Channel Availability Check condition: near perfect detection of each individual pulse because there is no wireless LAN traffic to interfere with detection. In fact, the graphs show that even 50% of loss of individual pulses does not affect the overall detection probability. In other words, Channel Availability Check detection is quite robust, regardless of the pulse count. However, the compliance criteria for this mode of detection are the same as for ISM mode detection.

The picture clarifies the degree of under-specification as well as over-specification of the current DFS compliance criteria: most of the detection space is not covered and, therefore, the behavior of a given implementation outside of the criteria is not assessed. As noted above, this presents a risk for the incumbents as well as for the wireless LAN industry.

Figure 9.7 shows the basis for improving the compliance criteria for radar detection. The rectangle indicates the range of parameter values that *together* determine the ability of a DFS detector to detect a variety of radar signatures under different conditions. The DFS compliance test criteria should be spread over a major part of that parameter space. In practice, the test "burden" need not be affected significantly: the test points need not be many and there is little need to test with many different pulse rates.

The dots mark 6 test points that together test a detector over its performance space, with different burst lengths and with different wireless LAN channel busy

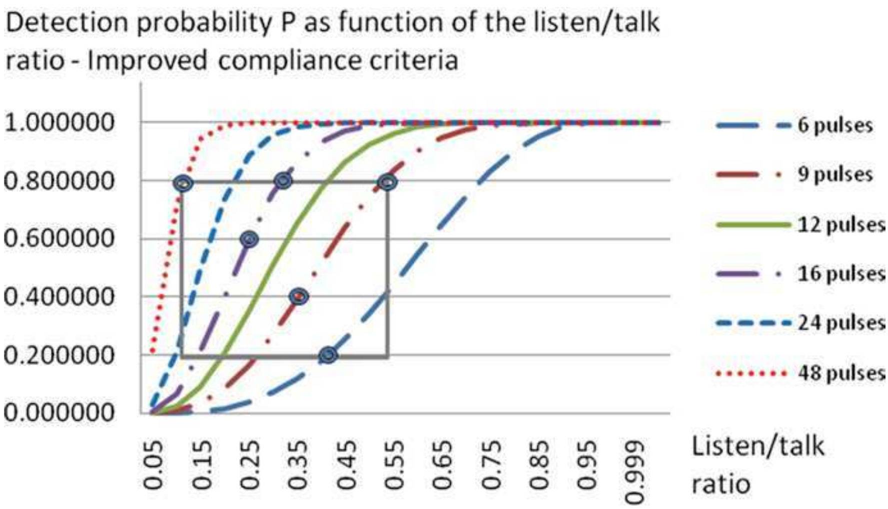

Fig. 9.7 Proposed DFS 2.0 validation points

ratios. A detector that shows this behavior will also perform according to these curves at the very high detection probabilities that are too time consuming and therefore too costly to test exhaustively. Tests to validate compliance can easily be automated, notably with assistance of testability features in the products to be tested, e.g. disabling the channel switch for and reporting the time, pulse count, and pulse spacing of the detected bursts.

9.2 Wireless LANs and 4G Technologies

9.2.1 Introduction

"4G" refers to a range of technologies that have been proposed as successors or alternatives to the third-generation cellular networks, which have become the dominant mobile access technology in the early years of the twenty-first century. The ITU-R has established "International Mobile Technologies" – IMT – as their label for cellular communications technologies. The more broadly used 3G term maps to IMT-2000, 4G maps to "IMT Advanced." Whereas 3G covered two technologies: Wideband CDMA and CDMA-2000, 4G will cover every contending technology that claims the ability to meet the IMT-Advanced requirements worked out by the ITU-R. Two key contenders are LTE – the long-term evolution of "3G" and WiMAX. LTE and WiMAX are becoming less distinguishable as time goes on.

WiMAX started life as IEEE 802.16 back in 2001 as a fixed point to multipoint system; it has seen various transformations since. The current standard is IEEE 802.16e, which was ratified in 2005; its successor is known as IEEE802.16m – which

is under development. WiMAX is the name of the trade organization that sponsors the development of products and test suites for these based on the IEEE802.16 standard. For ease of reference, the following will refer to WiMAX for technology; the IEEE standards will only be called out when relevant.

Initially, WiMAX[8] was designed to deliver IP broadband capacity at relatively long distances at acceptable cost. The use of Time Division Multiple Access (TDMA) operation allows low-cost designs and it assures IP compatibility by freeing the up/down capacity from the fixed up/down ratio associated with FDD. Long-distance operation was realized thanks to high gain antennas at the fixed clients. However, seeing a less than rosy market outlook for such systems, the WiMAX proponents switched to mobile instead of fixed access. Thanks to its TDMA basis, high performance IP services over WiMAX mobile seemed a no-brainer. Its IEEE counterpart is 802.16e. However, few looked at the inter-cell and intra-cell interference associated with TDMA operations. The "solution" was to synchronize the base stations and the base station sectors – an interesting challenge in a competitive business environment. In any case, the synchronization erodes much of the touted performance claims of TDD and reduces spectral efficiency.

Since the WiMAX community began to claim spectrum for its technology, the debate has raged about the relative spectrum efficiencies of the TDD and FDD.

From a spectrum sharing point of view, TDD/TDMA is the more interesting technology, since it may see use in both license exempt spectrum and licensed spectrum that is adjacent to license exempt spectrum. Therefore, the following will use WiMAX as example in the understanding that WiMAX too has FDD modes of operation.

WiMAX standards allow for operation in license exempt spectrum and some companies market it as a solution for customers who, for one reason or another, have no licensed spectrum. The overlap with wireless LANs occurs in the 2.5 GHz band and in the 5.4–5.8 GHz band. In the former case, the issue is adjacent band coupling only; in the second case, all variations of coupling and interference may obtain.

9.2.2 Adjacent Band Operation

Adjacent band operation between WiMAX and wireless LANs is possible in the 2.3 and 2.5 GHz regions. The wireless LANs operate between 2,400 and 2,483.5 MHz, WiMAX allocations in the lower 2 GHz band go up to 2,400 MHz; in the higher 2 GHz band, they go down to 2,500 MHz. Under the EU regulation, wireless LANs can use channels 1 through 13 – the latter being centered at 2,472 MHz and extending up to 2,480.5 MHz for OFDM and up to 2,483 MHz for CCK. In practice, OFDM is the dominant type of modulation. WiMAX operation above 2,500 MHz typically starts at 2,502.5 MHz provided the operator adheres to the applicable band edge mask. In case of an FDD deployment, the lower frequencies are

[8]This version of WiMAX corresponds to the IEEE 802.16d standard.

used for the uplink. This means that the separation in the lower 2 GHz band between the two systems is formally 0 MHz, whereas it is in the order of 15 MHz in the upper 2 GHz band.

Given the requirements for out-of-band emission levels for both systems allows one to assess the inter-system interference, statically and dynamically. The static Analysis looks at the link budget only; the dynamic analysis looks at interactions of the medium access protocols.

9.2.3 Link Budget Analysis

There are two different deployment cases to be addressed: (a) the outdoor WiMAX client and the indoor wireless LAN network, and (b) the co-located wireless LAN client and WiMAX client – as would be the case in a notebook or smartphone. In case (a), separation distances play a major role in the interference assessment; in case (b), the main factor is the separation in frequency and the transmitter mask/ receiver filter steepness (Fig. 9.8).

The transmitter mask used here is that for a Type HC system,[9] which has a minimum downlink rate of 8 Mb/s for a 7 MHz channel and is designed to be "highly co-existent." The transmitter mask is given by the figure below and the following data points: A = 2.8, B = 5.6, C = 7 and D = 14 MHz. A higher capacity variant such a Type B system offers twice the data rate of the HC type, but it needs 10 dB more SIR and, therefore, it is more sensitive to interference.

Case (a) Outdoor WiMAX client and indoor wireless LAN network
Many WiMAX deployments provide fixed services: directional antennas at the base station and clients are used to extend the range at higher bitrates to distances that allow WiMAX to be used as a DSL replacement in underserved, low-density areas.

The configuration is typically an outdoor WiMAX client with a medium gain antenna built into the package that is mounted outdoors on a wall. The wireless LAN

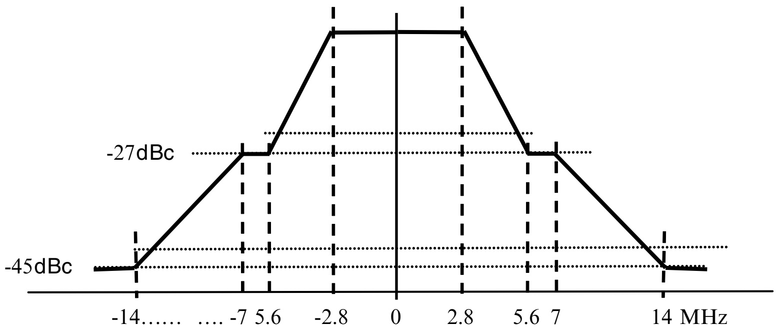

Fig. 9.8 WiMAX transmitter mask (HC system)

[9] See ETSI EN 301 021-2007.

network is represented by an Access Point at the other side of a wall and a distance of 2 m. A less worst case configuration would be that the Access Point is behind two walls and at a distance of 8 m. Assuming a wall loss of 12 dB and an antenna front to back ratio of 12 dB gives a 24 dB signal loss. On this basis, the respective coupling loss factors work out at 24 + 47 + 6 dB = 77 dB and 36 + 47 + 18 = 101 dB. Using the lower figure, we get the co-channel interference levels given in Table 9.4.

Since the WiMAX devices may be operating in the edge slots of their frequency bands, any frequency isolation comes from the distance of the wireless LAN's emission mask to the band edge. At the lower edge of the wireless LAN band, the two systems operate in adjacent channels. At the higher edge of that band, wireless LANs have to leave the upper 16.5 MHz unused (except in Japan, that case is ignored here).

With the WiMAX client output set to 37 dBm and the wireless LAN output set to 23 dB, the interference levels in case of a single wall separation work out as follows for the WiMAX to wireless LAN interference and vice versa (Table 9.5).

The other way around the picture is different, mostly because the wireless LAN radiated power output is much lower: 23 dBm instead of 37 dBm. In fact, since the technologies are comparable in many ways, the filter roll-off characteristics are likely to be similar (Table 9.6).

Table 9.4 Co-channel interference between WiMAX clients and wireless LAN devices

Radiated power	WiMAX: 37 dBm		Wireless LAN: 23 dBm	
Isolation	77 dB	101 dB	77 dB	101 dB
Interference level	–40 dBm	–64 dBm	–54 dBm	–78 dBm

Table 9.5 Outdoor WiMAX to indoor wireless LAN interference levels

WiMAX transmitter		Wireless LAN receiver		
Frequency offset (MHz)	Tx filter attenuation	Frequency offset (MHz)	Rx filter attenuation	Interference (dBm)
≤2.8	0	≤10	0	–40
5.6	27	11	12	–79
7.0	27	≤ 8	24	–91
14	45	20	36	–121
≥14	45	≥20	56	–121

Table 9.6 Indoor wireless LAN to outdoor WiMAX interference levels

Wireless LAN transmitter		WiMAX receiver		
Frequency offset (MHz)	Tx filter attenuation	Frequency offset (MHz)	Rx filter attenuation	Interference (dBm)
≤10	0	≤2.8	0	–54
11	22	5.6	12	–88
15	28	7.0	24	–106
20	32	14	36	–122
≥ 20	42	≥14	56	–152

The interference levels in above tables suggest that, for the indoor/outdoor case, a separation of 22 MHz between the centre frequencies of the two devices is necessary. Note that the interference thresholds of both systems have roughly the same value. This means that here is no margin available for detection of the system at a level where the detecting system will not cause interference.

Case (b) Co-located WiMAX client and wireless LAN client

The analysis of case (a) is easily adapted to the indoor, co-located arrangement of a WiMAX receiver on the table top next to the PC or PDA: the difference in attenuation is 30 dB and the interference levels are correspondingly higher (see Tables 9.7 and 9.8). The frequency separation must be increased to at least 35 MHz to protect the wireless LAN receiver from the more powerful WiMAX transmitter.

Case (c) Outdoor wireless LAN Access Points and WiMAX clients

This case may arise if frequencies used by outdoor wireless LAN deployments – e.g. mesh networks – are used by WiMAX systems. There are two sub-cases: the wireless LAN Access Point is close to the WiMAX base station and, secondly, the more likely sub-case of a WiMAX client near an wireless LAN Access Point. In terms of link budgets, this is a variation of the above co-located case: instead of a 1 m separation, the typical distance in this sub-case is 8 m. That gives 18 dB free space attenuation. Tables 9.9 and 9.10 show the impact. In practice, the frequency separation hardly required changes and more than 35 MHz separation is needed.

Table 9.7 Co-located-WiMAX to wireless LAN interference levels

WiMAX transmitter		Wireless LAN receiver		
Frequency offset (MHz)	Tx filter attenuation	Frequency offset (MHz)	Rx filter attenuation	Interference (dBm)
≤2.8	0	≤10	0	−10
5.6	27	11	12	−49
7.0	27	15	24	−61
14	45	20	36	−91
≥14	45	≥20	56	−111

Table 9.8 Co-located-wireless LAN to WiMAX interference levels

Wireless LAN transmitter		WiMAX receiver		
Frequency offset (MHz)	Tx filter attenuation	Frequency offset (MHz)	Rx filter attenuation	Interference (dBm)
≤10	0	1	0	−24
11	22	5	12	−58
15	28	8	24	−76
20	32	10	36	−92
≥20	42	≥20	56	−132

Table 9.9 WiMAX to wireless LAN interference, outdoor

WiMAX transmitter		Wireless LAN receiver		
Frequency offset (MHz)	Tx filter attenuation	Frequency offset (MHz)	Rx filter attenuation	Interference (dBm)
≤2.8	0	≤10	0	−28
5.6	27	11	12	−67
7.0	27	15	24	−79
14	45	20	36	−109
≥14	45	≥20	56	−129

Table 9.10 Outdoor – wireless LAN to WiMAX interference levels

Wireless LAN transmitter		WiMAX receiver		
Frequency offset (MHz)	Tx filter attenuation	Frequency offset (MHz)	Rx filter attenuation	Interference (dBm)
≤10	0	1	0	−42
11	22	5	12	−76
15	28	8	24	−94
20	32	10	36	−110
≥20	42	≥20	56	−150

9.2.3.1 Dynamic Channel Selection for Adjacent Band Operation

Although co-channel is not an issue because of regulatory constraints, adjacent band operation includes the possibility of adjacent channel operation. Presumably, one could apply dynamic channel selection to avoid this. Since the WiMAX system is likely to be deployed by a service provider who selects channel settings for his equipment, so as to optimize the performance of his network, it is the wireless LAN which would have to adjust. Such adjustment requires detecting the WiMAX signal at a level near or below the noise floor of the wireless LAN receiver in order to achieve a frequency separation of at least 30 MHz.

Note: Because of the asymmetry in power levels – which is not compensated by the difference in bandwidth, WiMAX seems to have a distinct advantage: the wireless LAN receiver suffers interference from the WiMAX transmitter, whereas the WiMAX receiver does not necessarily see the wireless LAN transmitter. On the other hand, the wireless LAN protocol allows recovery from the intermittent errors caused by the WiMAX transmitter, whereas the reverse may not be true. See also Sect. 9.2.4.

From a channel selection point of view, the stronger WiMAX signal should be a boon. A typical WiMAX receiver with a 5 MHz bandwidth and 4 dB noise figure has a receiver sensitivity of −103 dBm. A protection margin of −6 dB I/N suggests that to realize its full capacity, the interference into the WiMAX receiver should not exceed −109 dBm. For a 16.5 MHz wide wireless LAN receiver, that figure is −98 dBm. In other words, the wireless LAN minimum signal level is 11 dB over that of WiMAX receiver. If the WiMAX signal is just below the detection threshold

of the wireless LAN, the WiMAX SIR is effectively 11 dB. Because the wireless LAN receiver will need a signal of at least 6 dB above its noise floor to reliably detect the WiMAX signal, the SIR margin is reduced to some 5 dB. Implementation considerations like calibration uncertainties may further reduce this value. This margin can only be increased by requiring the wireless LAN receiver to detect the WiMAX uplink training sequences. In practice, this would mean that wireless LANs would have to be designed to detect WiMAX systems. Given the independent development of these technologies, this is not a realistic approach.

To conclude: co-located or nearby operation of WiMAX systems and wireless LANs require a frequency separation of at least 40 MHz for mobile devices and at least 35 MHz for fixed WiMAX clients. Automatic frequency adjustment, based on the detection of WiMAX signals by the wireless LAN system to avoid interfering with a WiMAX client, is not possible.

9.2.4 Protocol Interaction Analysis

The preceding static analysis raises the question of interaction in case the WiMAX client and the wireless LAN operate at less than the minimally desirable frequency separation. This problem can become acute if WiMAX systems are put into operation in bands that are heavily used by wireless LAN systems. As the preceding sections show, interference in co-channel operation is high and even in first adjacent channel cases it is sufficient to cause data loss. Therefore, it is relevant to look at the protocol of both systems to see what the impact of the interference is.

The interference situation in general is like described above in the static analysis: wireless LAN transmissions may interfere with the WiMAX downlink, WiMAX uplink transmissions may interfere with the wireless LAN “downlink” transmissions. For the sake of simplicity, the following only considers the case of a WiMAX client in the vicinity of a wireless LAN client, but the arguments presented can easily be generalized include the case of a WiMAX client near – or co-located with – an wireless LAN Access Point.

9.2.4.1 Wireless LAN Traffic Patterns

The basic wireless LAN protocol elements are shown in Fig. 9.9, which covers a period of roughly 3 ms. A frame may be a beacon which carries traffic management information – e.g. sleep/wake messaging – and, therefore, loss of beacons may interfere with the scheduling of data traffic between an Access Point and the client that suffers interference. Loss of data occurs on the downlink only when the WiMAX client transmission overlaps with the reception of a wireless LAN frame. Frame duration varies e.g. from a very short 210.2 μs for a 64 byte frame sent at 54 Mb/s to a long 1.35 ms for a 1,000 byte frame sent at 6 Mb/s. Wireless LAN data rates are adapted to the link conditions: a weak link will cause a lower data rate to be used.

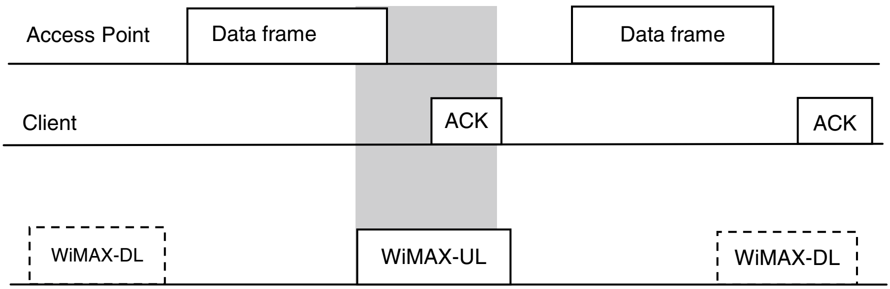

Fig. 9.9 WiMAX–wireless LAN interference scenario

Thus, for an wireless LAN client that is far removed from its Access point, interference carries a double burden: loss of data and – because of the long transmission time — a high probability of such loss. The following example clarifies this: assume a remote wireless LAN client is performing a file transfer; frames of 512 bytes are transmitted at intervals determined by the traffic level in the Access Point's network. If the Access Point is not very busy, the remote client may get service roughly 50% of the time or more. Since the client is remote, the data rate is 6 Mb/s and frame transmission time is .702 ms. Applying the assumed 50% rate of service, the wireless LAN client would receive a frame every 1.4 ms or a maximum of 700 times per second. A busy WiMAX client may transmit at a similar rate and destroy a large fraction of wireless LAN frames.

A side effect of rate adaptation is that the robustness of the wireless LAN link against interference varies little with distance or data rate: the rate adaptation will seek the highest possible rate that the link conditions – primarily SNR and/or SIR – will sustain. Such data loss, could cause a lower data rate to be used, will be increased because of the increased interference window.

9.2.4.2 WiMAX Traffic Patterns

The WiMAX protocol is a point to multipoint protocol that makes use of frames, the duration of which may vary per installation. The shortest frame duration is 2.5 ms. As shown, the WiMAX downlink – which may be longer than shown – occurs at the frame rate and in between the WiMAX uplink occurs. The downlink starts with a frame header, which contains a downlink map and an uplink map. The WiMAX clients use these to determine their receive and transmit times. Header loss at a client forces that client into a recovery mode. The recovery process relies on the frame header and, therefore, interference while the frame header is being received is particularly damaging. A related issue is that much of the downlink transmissions are broadcasts and interference during these broadcasts may cause loss of data.

Interactions of Wireless LAN Clients and Proximate WiMAX Clients

1. Wireless LAN as victim of a nearby WiMAX client
   Given the WiMAX frame duration of 2.5 ms, the frame rate is 400 frames per second. The activity of the WiMAX client depends on the service being provided. Assuming the WiMAX data rate is 240 kbps, a file transfer or video call running at 40 kbps would occupy roughly a 6th of the frame time. This would leave most of the wireless LAN's downlink capacity untouched.

   As shown above, the first data frame would be affected, the Ack would be sent and the Access Point would reschedule the same frame for transmission. If the WiMAX uplink transmission interfered with a beacon, the effect would depend on the details of the situation. If the wireless LAN client is operating normally, loss of a single beacon will have no impact. If the wireless LAN is coming out of a sleep period, it may expect a beacon at a certain point in time. This process repeats at the sleep/wake rhythm of the wireless LAN device – typically in the order of many seconds. If such an expected beacon is missed, the wireless LAN client will wait for the next beacon. Thus, the impact of WiMAX client interference on the wireless LAN sleep/wake process is not significant.
2. WiMAX client as victim of a nearby wireless LAN client
   Since the assumed wireless LAN frame rate is as high as 700 frames per second overlap, and therefore interference is unavoidable. At full load, the WiMAX service rate is 1 every 2.5 ms. Assuming the WiMAX client is receiving a video broadcast with a data rate of 40 kbps, the frame size is roughly 200 bits plus header information sent at every 2nd frame. If the WiMAX transmission rate is in the order of 240 kbps, the duration of the downlink video frame is approximately .8 ms or roughly a third of the frame time. With the wireless LAN frames coming at 1.4 ms intervals, interference is assured on every second frame.

   In case of a voice service instead of a video service, the picture is less bleak: at 8 kbps, the frame rate could be dropped to 1 in 3 and the duration would drop to .5 ms. However, the tight frame spacing of the wireless LAN downlink would cause a high frame loss at the WiMAX client: if the wireless LAN is active 50% of the time, it occupies 1.25 ms per frame in blocks of .7 ms. The probability of a WiMAX transmission suffering interference is 1/1.7 or 59.8%. For all practical purposes, this loss rate is far too high. If it is reduced by a factor 12, reasonable voice and video service would be feasible. Such a reduction is possible only if the two devices are co-located and controlled by the same entity.

9.2.4.3 The Potential Benefit of FDD

Applying FDD mode to the WiMAX system allows the removal of the source of performance loss in the downlink – wireless LAN interference. The asymmetry introduced is significant: the wireless LAN will not affect the WiMAX downlink, but the WiMAX uplink will affect the wireless LAN downlink – and vice versa,

depending on the relative frequency distances. Only if the spectrum is partially licensed or otherwise controlled, such that wireless LANs can be kept away from WiMAX downlink frequencies, FDD would actually provide a significant benefit. Conversely, wireless LANs will benefit from not having WiMAX uplink transmitters in an adjacent channel.

9.2.4.4 Contention-Free Wireless LANs

As noted above, wireless LAN clients and WiMAX clients could be operated in the same frequency band and this leads to potentially serious interference levels, even if co-channel operation is completely avoided.

The preceding dynamic analysis makes it clear that the protocols of WiMAX and wireless LANs do not allow interference to be controlled or reduced to acceptable levels. This is a direct consequence of the difference in medium access control procedures: whereas one is nearly random, the other is strictly ordered. Clearly, turning WiMAX into a semi-random access protocol, like the wireless LAN protocol, would not solve anything: not only would the performance of the WiMAX system drop significantly, but coordination between the two would not be feasible. The reason is that CSMA requires all participating stations to be in detection range.

The alternative is to change the wireless LAN protocol and use the contention free mode, with the Access Point acting as the Point Coordination Function (PCF). Presumably, the PCF could coordinate the wireless LAN transmissions for a given client with the transmissions from the WiMAX client, such that interference would be avoided. This is not the case without major changes to the WiMAX protocol: the coordination has to work both ways: the wireless LAN has to adjust to the WiMAX downlink as well as to the WiMAX uplink. Since this cannot be done on a frame-by-frame basis, the WiMAX base station would have to fix transmissions to and from a given client for long periods of time. This would interfere with the scheduling of the WiMAX base station of the servicing of its clients and lead to unacceptable reductions of throughput and/or response time. This reasoning begs the question of how coordination is to be achieved: as the static analysis shows, there are cases in which the two types of client can interfere with each other without being able to detect each other reliably. Therefore, coordination would have to be implemented with the aid of out-of-band signalling. Proposals in that direction have been made in the IEEE 802.16 Working Group, but this has not been taken up by its 802.11 sister group and implementations have not been put on the market.

9.2.5 Summary

The preceding sections provide an analysis of some examples of if and how a TDD WiMAX system would survive operating at frequencies that are adjacent to or overlap with frequencies used by wireless LANs. WiMAX can be generalized to stand

for any 4G technology and, similarly, "wireless LANs" can be generalized to stand for any unlicensed radio technology. The conclusions are negative, except for the case in which the 4G system operates in FDD mode, with its downlink in spectrum that is adjacent to the license exempt spectrum. In that case, the 4G downlink frequencies would be farther away from the license exempt band.

9.3 License Exempt Systems in the TV White Space Bands

TV White Space spectrum is spectrum made available for a radio communication applications on a non-interfering/non-protected basis. Availability varies with time and geographical location due to the need to protect broadcasting – the established primary users.

9.3.1 TV White Spectrum Applications

Two types of candidate license exempt devices are being considered for deployment in the UHF bands:

– Personal/portable devices, such as wireless LAN cards in laptop computers, smart phones, tablet and other commodity gear which operates with, for example, a maximum power output of 100 mW.
– Fixed/access devices that are generally operated from a fixed location and may be used to provide a commercial service, such as wireless broadband Internet access. These devices could operate with a transmitter output power of up to 1 W.

The applications foreseen for white space spectrum include wireless LAN networking and long-range bridging, community mesh networks, in-home networking and home automation. Each TV channel would support data rates that are comparable with those of current 3G mobile services. White space devices may have to aggregate any available channels to achieve more throughput. On the other hand, for some applications, the throughput of a TV channel may be far more than is needed. This suggests a potential trade-off between bandwidth and range: high throughput applications will use the full bandwidth of one or more channels, whereas low throughput applications such as home automation could opt for a narrowband approach to gain the energy density necessary for extended coverage and wall penetration. As a consequence, dissimilar types of TV White Space devices might find themselves competing for the same TV channel. See Sects. 2.3.1 and 2.3.2 for the regulatory background and Sect. 4.2.4 for more background on the use of a geo-location data base to facilitate sharing of this spectrum.

The following sections will only consider the TVWS sharing regime in the US.

9.3.2 Sharing TV White Space Spectrum

9.3.2.1 The Regulatory Requirements

As indicated above, the FCC saw the potential of the TVWS spectrum more in terms of creating opportunities for private initiative to address the needs of society. Hence, the referenced Second Memorandum, Opinion and Order on TV band devices (TVBD) of September 2010.

The FCC's Part 15 rules in paragraphs 700 through 717 define the rules under which TV Band Devices may access the TVWS spectrum. That spectrum consists of a number of frequency bands and there are rules for its use that are related to the operation of a device: Mode 1 and Mode 2 and fixed or mobile. A Mode 1 device is not allowed to operate in these frequencies without being enabled by a device acting as Master: either a fixed device or a mobile Mode 2 device – these roles require that the device is capable of selecting an operating channel, using the methods defined by this part of the FCC's rules (Part 15 paragraph 711).

Table 9.11 lists how role and spectrum usage are related.

The RF parameters and requirements are as follows (Table 9.12).

In addition, an emission mask is defined around the band 608–614 MHz for the protection of the radio astronomy service. Spurious emissions must be limited to the general level of such emissions.

The basic interference mitigation mechanisms that these devices are required to implement are given as follows:

(a) Use geo-location and access a spectrum management database or perform spectrum sensing in order to determine if a channel is free.
(b) Avoid operation within a given separation distance outside the protection contour for a given channel. This separation distance depends on antenna height and channel separation.

In addition, there are a number of exclusion zones related to TV translator receive sites and cable head-ends, related to the fixed Broadcast Auxiliary Service, around

Table 9.11 TVWS frequencies and mode of use

Operating mode	Frequency band (MHz)				
Fixed–fixed	54–60	76–88	174–216	470–608	614–698
Fixed–mobile				512–608	614–698
Mobile–Mode 2				512–608	614–698
Mobile–Mode 1				512–608	614–698

Table 9.12 TV White space – RF parameters

Operating mode	Allowed Tx output	Allowed EIRP	ACLR[a]
Fixed	<1 W	<4 W	72.85 dB
Mobile	100 mW	100 mW	72.85 dB

[a]Relative to the total in-band power, measured in 100 kHz

Table 9.13 TV White Space contour protection distances

	Required distance to protection contour (km)	
TVBD antenna height (m)	Co-channel	Adjacent channel
0–3	6.0	0.1
3–10	8.0	0.1
10–30	14.4	0.74

Table 9.14 Protection contour field strength levels

Channels	Field strength (dBu) for F(50,50) curves	In dBm	Field strength, in dBu, for F(50,90) curves	In dBm
2–6	47	–60	28	–79
6–13	56	–51	36	–71
14–69	64–20 log (615/channel frequency)	–46 to –43	41–20 log (615/channel frequency)	–69 to –66

metropolitan areas and radio astronomy sites and near the borders of Canada and Mexico. Although these exclusions may seem to cover most of the continental US, reality is less severe.

The required separation distances for co-channel and adjacent channel operation of a TV band device depend on antenna height (Table 9.13).

The crux of this rulemaking lies in the TV Bands Database.[10] The purpose of the TV Band Database is given as:

(a) To provide a requesting TV Band Device with the available channels at the latter's location.
(b) To register protected locations and channels in addition to information available in existing licensing databases, sites using wireless microphones, etc.
(c) To register the location and identification of fixed TV Band Devices as well as sites.

In addition, there is long list of items and criteria related to the use, content and updating of a TV Band Database. Such a database may be a commercial operation; the FCC has the power to review and adjust the database access fees.

The key content of the database – in addition to the known geographical location constraints on TV Band Device operation – are the protection contours for a given TV station and its operating channel(s). These contours are established by existing TV regulations – given in CFR73.699 with regard to interference control among TV operators. Thus, a protection contour determines the TV signal level needed to achieve adequate picture quality with an off-the-shelf receiver.

The field strength at the protection contour depends on the type of TV (Table 9.14).

These figures are for the power measured over the whole channel (6 MHz) and reflect signal levels that are adequate for viewing. Note the 23 dB difference between analog and digital TV contour levels: digital TV is much more sensitive to interference than analog TV.

[10]For a general perspective on spectrum sharing data bases see Sect. 9.4.2.

Table 9.15 Main DTV Receiver parameters

Thermal noise (dBm)	106.2
Antenna Gain (dB)	10
Downlead line loss (dB)	4
System noise figure (dB)	7
Required Carrier/Noise ratio (dB)	15

9.3.2.2 Analysis

1. Protection at the Protection Contour
   The basis for the protection contours applied by the FCC is given in FCC, Bulletin OET-69, "Longley-Rice Methodology for Evaluating TV Coverage and Interference," February 06, 2004. For the UHF band, it uses the following figures (Table 9.15).
   In effect, the net antenna gain is 6 dB – in the same range as the system noise figure. This suggests that a minimum SIR of 16 dB is needed to assure decent TV picture quality. This figure is relevant for the main lobe of the TV antenna. However, an analysis of intra-service interference[11] in ATSC networks shows that in field tests a margin of 20 dB is needed at the signal processing input of a DTV receiver. Using this figure as the required SIR at the protection contour at channel 21–615 MHz – gives a maximum interference signal of −66 – 20 dBm = −86 dBm for the front lobe of the TV antenna and −76 dBm at the back lobe of the TV antenna. Adding a 10 dB fading margin gives −96 and −86 dBm respectively.
2. Separation distances for fixed TV Band Devices
   In order to keep the interference signal at the above levels, the separation distance has to provide significant attenuation.
   For the Mode 1 and Mode 2 devices, which are portable and/or nomadic, the pathloss profile with three breakpoint towards at DTV antenna can be given as [2/100 m; 2.5/1,000 m; 3.3/infinity]. For a 100 mW transmitter, the required separation distances are 2,630 and 1,309 m respectively. Tests[12] performed by the FCC included anecdotal results that confirm that a 5 mW device operating at ground level at a distance of 360 m will cause co-channel interference into a digital TV receiver with an antenna at a height of 10.3 m. This fits with the results obtained using the three breakpoint pathloss model for mobile TV band devices.
   For a fixed TV band device, the pathloss profile would be different: the higher antenna position gives a lower attenuation. Allowing for some vegetation in path gives a single breakpoint path with an exponent of 2.5 after 1,000 m up to the local horizon; beyond that, a 30 dB/octave attenuation is assumed in these examples to take into account the rapidly increasing attenuation beyond the horizon. Table 9.16 summarizes the above and adds the FCC's protection distances.

[11] See Bendov [22].

[12] See FCC/OET 08-TR-1005.

Table 9.16 TV band device separation distances

TV band device	TV band device location		FCC requirement (m)
	Inside contour (m)	Outside contour (m)	
Mobile/nomadic	6,036	2,652	6,000
Fixed	50,451	20,625	14,400

In practice, digital TV receivers may well be able to operate at 14 dB lower receive power – this equates to separation distances of 8,386 for mobile TV band devices and 55,318 m for fixed TV band devices with an antenna at 10 m height. These distances include a 10 dB fade margin for the wanted signal and, therefore, actual distances could be larger. The separation distance for the fixed TV band devices is larger than the FCC's protection margins; given the variation in propagation conditions, interference complaints are to be expected once TV band devices are entering general use.

9.3.2.3 Protection Based on Sensing DTV Transmitters

1. Sensing by Fixed TV band Devices
   The separation distances given above for outside the protection contour can be used to assess the required DTV sensing threshold for TV band devices. The wanted signal at the contour is −56 dBm. For a 10 KW transmitter, this corresponds to a distance of 66,354 m using the same pathloss model as above for the Fixed TV band device and the TV receiver. The 4 W Fixed TV band device would have to be at a distance of 52,707 m in order to assure the 20 dB SIR requirement at the DTV receiver.

   Ignoring the extended TV coverage area, the worst case detection scenario is where the TV band device is located along the same line as the DTV transmitter relative to DTV receiver. The distance between TV transmitter and TV band device is then 85,440 m. This exceeds the line of sight distance to the local horizon for both sides.

   Applying the above pathloss model with 30 dB/octave beyond the horizon gives insight into the required detection threshold for the DTV signal. Such a pathloss model is a 3 breakpoint profile with the following parameters: [2/1,000; 2.5/10,000; 2.8/50,000; 10/infinity]. Assuming that the detector requires a signal of 6 dB above the thermal noise floor gives a detection distance of 76,185 m – which is not adequate to protect the TV receiver. In other words, energy detection only will not work. Instead, the detector has to be signal specific and leverage redundancy, so that detection below the noise floor becomes possible. In the example used here, lowering the threshold by 6 to −106 dBm achieves the required detection distance: but only just. However, since the signal to be detected is subject to shadowing, at least another 10 dB are needed to assure a reasonable degree of protection at the protection contour. Thus, the detection threshold has to be set at −116 dBm or lower for a 6 MHz signal. This equates to −122 dBm/MHz and gives a detection distance of 110,121 m. These results are close to the

detection threshold defined in the IEEE 802.22 standard for Wireless Regional Area Networks.

2. Sensing by mobile TV band Devices
   Since mobile and nomadic TV band device operate at low level, the propagation losses are higher. This can be modeled by changing the above pathloss parameters to [2/1,000;2.5/6,000;3,3/18,000/10/infinity]. The resulting detection distance for energy detection is below the local "flat earth" horizon: 32,246 m. Lowering the threshold to −122 dBm/MHz gives a detection distance of 46,609 m, which is still below the required detection distance of 69,006 m. This result indicates that mobile devices – or in general devices operating at low levels above ground, will not be able to detect protected service receivers at adequate distances. The main reason is the high attenuation of the last segment of the path-loss profile between DTV transmitter and the detecting devices.

9.3.2.4 General Conclusions

Table 9.17 summarizes the results of the above analysis.

A number of conclusions can be derived from this table:

1. The directional antenna assumed to be used for the DTV receiver (with a front to back ratio of 10 dB) makes a large difference in the interference budget: it reduces the required separation distance by a factor 3 or more.
2. The antenna height of the Fixed TV band device adds considerably to its unfavorable interference potential, but that is countered to some extent by the better detection distance for DTV signals.
3. Energy detection will not provide enough detection range to protect DTV broadcast receivers.
4. Dedicated signal detect can achieve the necessary detection range, but only if propagation conditions are favorable.

Taking a step back from the details of this analysis suggests that energy detection will never be adequate to assure adequate protection distances. Therefore, all attempts to devise spectrum sharing schemes – including the highly successful CSMA/CA scheme of the IEEE 802.11 standard – are bound to fail in general; they will only work in special cases. An example of such a special case is the highly asymmetrical case of radar-wireless LAN sharing, described in Sect. 9.1 above.

Table 9.17 Summary of typical separation and detection distances needed to protect DTV from TV band devices

Case	Fixed TVBD	Mobile TVDB
Separation distance inside the protection contour (m)	43,437	6,036
Separation distance outside the protection contour (m)	19,086	2,652
Required detection distance, outside protection contour (m)	85,440	69,006
Actual detection distance, energy detection (m)	76,185	32,246
Actual detection distance, DTV detection (m)	110,121	46,609

9.4 Cognitive Radio

Ever since Mitola's seminal work on Cognitive Radio,[13] this concept has captured the imagination of spectrum regulators, engineers and, not to forget, pundits. Cognitive Radio technology is a concept that marries information technology, communications technology and radio technology, with the aim of allowing radio devices to determine the spectrum to use at a given time and location instead of being restricted to fixed regulatory spectrum allocations.

Great advances in spectrum availability and spectrum efficiency were predicted by many, and various frequency bands have been proposed as suitable for cognitive radio technology. As noted in the preceding sections, TV White Space spectrum holds promises as well as challenges and, therefore, it provides a potential "training ground" for Cognitive Radio technology. If Cognitive Radio technology proves practical in this context, it will see broader use.

The primary tools in the Cognitive Radio tool kit are:

(a) Spectrum sensing
(b) Geo-location and spectrum database
(c) Spectrum Management Beacons

Each of these has been studied extensively in different contexts and with different levels of detail.[14] A common element in the downsides to each of these three is the variability of radio propagation. The following sections address this.

9.4.1 Spectrum Sensing

Spectrum sensing as tool for coordinating frequency use is not new; it proved a suitable tool in the case of wireless LANs sharing the 5 GHz band with radar system. Spectrum sensing also played a major role in the work of Mitola and many others – notably the assumption of its feasibility.

In vertical spectrum sharing[15] cases, spectrum sensing is used for detection of the primary spectrum users. Examples are given in the preceding sections of this chapter: Wireless LANs in the 5 GHz band sensing radars and license exempt devices in the UHF bands sensing TV transmitters.

Spectrum sensing in horizontal spectrum sharing is used for coordination among devices operating in the same spectrum: sensing supports both mutual avoidance of co-channel operation – like it does in the case of vertical sharing – as well as coordinated use of the same spectrum, for example in the CSMA/CA protocol of IEEE 802.11.[16]

---

[13] See Mitola [97].

[14] A good example of an extensive and detailed analysis is (draft) ECC Report 159.

[15] See also Sect. 4.1.1.

[16] See Sect. 6.2.3.

These examples testify to the difficulties of reliable spectrum sensing: in the case of wireless LAN/Radar sharing, the sensing works well – thanks to the very high power of the radar emissions and the distinctive radar pulse patterns. The high power of the radar emissions assures detection over a wide range of propagation conditions. However, detectability depends also on adequate redundancy in the radar pulse patterns.

In the case of the TV White Space sharing, the large asymmetry between the TV transmitter output and the TV band devices is not adequate to assure detection under all circumstances – notably, the mobile devices cannot detect the TV transmitter signal whenever this is necessary to protect the TV receivers. This is due largely to the unpredictability of propagation conditions.

In the case of sharing between devices and systems with the same status, the efficacy of detection is further reduced because propagation conditions play a crucial role when the transmitted power levels are comparable. This observation holds for any power level and any realistic propagation scenario. Only in free space – well outside geo-stationary orbit – the propagation conditions are sufficiently homogenous to make reliable detection of peer systems possible.

The shortcomings of spectrum sensing have been widely recognized and analyzed,[17] and various ways of addressing them have been proposed, including cooperative sensing. However, the major limiting factor is the fact that, for a given detector, the number of observations required for reliable detection approaches infinity as the SNR drops. This is called the "SNR Wall."[18] The position of this SNR Wall is determined by the operating SNR, the noise uncertainty level of the detector and the coherence time of the channel. The latter adds an environment factor, which varies in time and space and which limits the detection robustness in a way that cannot be captured in rulemaking, whether regulatory or voluntary. Therefore, safety margins are required. Disregarding channel coherence limits, the SNR wall for a simple energy detector with a 1 dB noise uncertainty is −3.3 dB; for a coherent detector it is 10 dB better: −13.3 dB. Given the need for higher protection margins in practice – e.g. −22 dB in case of TV transmitters in the UHF bands – this is not adequate. Cooperative sensing has been proposed as a means to circumvent this limitation of individual sensing, but cooperation has to be such that the signals from other devices do not increase the detection uncertainty for a given detector; this leads to the need for a central "detection manager." In any case, lowering of the SNR through integration of sensing results wall requires listening periods that interfere with system operation.

Although cooperation can be designed into a system, the means of coordination will tend to be system unique and, therefore, the practical benefit of cooperative detection is limited: non-cooperative systems will create the noise uncertainty that moves the SNR wall closer (= higher). Therefore, even if spectrum sensing is designed and implemented at system's level, its usefulness is limited.

---

[17] See Sahai [110].

[18] See Tandra and Sahai [120].

9.4.2 Geo-Location and Spectrum Databases

Geo-location refers to the capability of a device to determine its position geographically. Given adequate accuracy of this function and a database of available and/or blocked frequencies per location, a cognitive device could make an autonomous determination what, if any, frequencies are available for use at its current location. This begs the question about condition at the intended receiver; that subject is not further addressed here. Presumably, geo-location could address the issues spectrum sensing (detection performance) noted above, because it is not – at least not in principle – limited by the SNR wall.

In its simplest form, the above takes place at the time a device or system is installed: with the aid of the coordinates of the location, the locally available spectrum is determined – e.g. with the aid of a public database and the transmitter of the devices is so programmed. Depending on the type of system envisaged, such a device can be used as enabler for other devices – this is addressed in the next section.

Geo-location has a certain inherent accuracy; this puts a lower limit on the resolution of the spectrum management database. The upper limit of that resolution depends on the scale of the propagation effects to be taken into account: on the plains of the mid-western US, the "pixel" size can be much larger than in Manhattan. If a 100 × 100 m pixel size is required, 8,750,000 pixels would be needed for Manhattan only and for each pixel a raft of information would have to be stored – not just the available frequencies, but also the allowable transmit power levels, time of day variants, validity qualifiers, etc.

Clearly, the creation and maintenance of such a database is a major undertaking. Even disregarding the scale and efforts related to such a spectrum management database, some issues remain visible. First of all, populating the database on the basis of surveys will be very time consuming if not impossible. Therefore, the database will have to "grow" with usage. Such a growth strategy raises many questions, for example:

- Accuracy: although the scale of the pixels can – and must – be adapted to the diversity of the environment being captured, the question of accuracy remains. What is the basis for storing certain values in certain pixels?
- Timeliness: although data content can be updated automatically – what is the frequency with which database is updated?
- Control and ownership: Given that such a database may have to be a commercial operation, the question arises who controls the content of the database in the longer term and guards its accuracy? Is the data publicly owned and is only the operation commercial or is the database operator to be considered the owner of the data?
- Access Control: who gets permission to effect updates and/or corrections? What are the implications of the data base being hacked?

The last three questions refer to operational issues, but the first one – accuracy – is an essentially technical one. Given the variability of the RF environment, it

matters at which point within a pixel measurements (= sensing data) are recorded and – because of the large impact of the uncertainty concerning the actual "noise level," it matters if there are other users on the channel at the time. In other words, the database creation is subject to the same issues as spectrum sensing by autonomous devices. Updating its content with data obtained after or during deployment of secondary devices may take a long time, during which the primary users run the risk of interference and the secondary users run the risk of spectrum not being available. Note that if the latter is the case, improvement is unlikely: if the database is too restrictive, the spectrum that the restrictions apply to may never be updated.

These and many other issues related to geo-location based spectrum sharing are well beyond the scope of this book.

9.4.3 Spectrum Management Beacons

As mentioned in the preceding section, secondary devices in a vertical sharing regime may be authorized to enable the use of some spectrum band by other devices. The FCC's Part 15 rules contain a definition of "master" and "client" roles for such devices.[19] Devices in the Master role are allowed to establish a connection or network involving other devices in the client role. The latter role forbids establishing connections without being "enabled" to do so by a device in Master role.

The Master/client role scheme, applied to secondary spectrum users, allows fixed devices in known locations and with verified interference profiles, to control other devices such that the protection of the primary users is assured. This principle has become embedded in the IEEE 802.11 standard under the heading of dynamic spectrum enablement or DSE.[20] It allows devices compliant with that standard to act as either Master or Client and so comply to the FCC's requirement that certain primary users are to be protected from interference, notwithstanding the 20 W maximum EIRP that fixed secondary systems are allowed to emit.

The DSE scheme may be assumed to operate reasonably well in practice – given the exclusion zones around the areas to be protected and the fact that the main type of use of the 3,650 MHz band is for point to multipoint systems. The base stations of these systems are the natural enablers for their clients.

However, extending the DSE scheme to a full blown beaconing system that is able to serve a wide variety of secondary users is another matter. Whereas a database system for spectrum sharing can be placed anywhere (it needs a little more than computers and an internet connection), a beaconing system with adequate utility would require large-scale deployment of beacon transmitters. That by itself is a major reason for the commercial lack of interest in such beaconing-based enablement schemes.

---

[19] See 47CFR15.201.

[20] This extension of the standard was known as 11y; it has been incorporated in the 2011 version of the standard.

However, beacons may be very practical for the purpose of protecting localized deployments of primary spectrum devices. The main problem here is to assure secondary users check out the beaconing channel prior to actual spectrum use. A problem with such a scheme is that some RF channel – close to the frequency band to be protected – has to be set aside for such disablement beaconing.

9.4.4 Summary

Cognitive radio is a concept whose time has come. As with all new concepts and technologies, it takes time for the market to embrace the concept and for applications to emerge. DFS – the wireless LAN-radar sharing scheme – proves that, at least in some cases, spectrum sensing will provide an adequate basis for the protection of the primary spectrum users. The initial work on a spectrum sharing database to support geo-location based sharing shows promise, but the market has yet to show it has adopted this means of spectrum sharing. Beaconing shows more promise as an embedded function of a particular secondary spectrum usage system rather than as a generic, widespread service that could become as ubiquitous as either of the other two methods. Disablement beacons promise to be more economical, but the required standards and rulemaking remain to be developed.

Given this state of cognitive radio technology, further development of DFS-like schemes, in combination with spectrum management databases, are a very promising direction for further development of this technology.

Chapter 10
Radio Resource Management

Previous chapters discuss the possible causes of degradation in the performance of commodity wireless technologies due to the properties of the technologies themselves, as well as due to the interactions between these technologies if they are operated in the same or adjacent RF spectrum. This chapter describes the use of Radio Resource Management as a means to optimize the performance of wireless networks that use these commodity technologies. These networks include wireless LAN-based networks, but may well include femtocells in the near future.

10.1 Introduction

10.1.1 Overview

Poor wireless network performance due to radio interference is among the top customer complaints for wireless network deployments. With the presence of interference, users experience degradation in wireless network performance, such as a slow communication and excessive jitter in voice applications. In the worst scenarios, wireless connections can be interrupted altogether. Much effort has been invested in addressing the issue, both in academic research and industrial implementations. The rapidly increasing use of wireless technologies has led to many and denser wireless network deployments than before, and this makes the issue even more challenging. Notably, denser deployments often introduce many more interference sources. Advances in wireless network technologies, such as MIMO, suggest that many applications over wireless network are now possible, which were not feasible before. Users have come to expect a much higher level of performance. Applications involving continuous media streams, such as embedded videos in internet applications, become more and more popular these days. This trend raises the overall traffic loads carried over wireless networks by one order of magnitude or more – which greatly aggravates potential performance and interference issues.

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_10,

Device manufactures have invested significantly in improving the performance of wireless network devices, by designing advanced modulation and coding schemes such as OFDM and MIMO. These advances in wireless technologies enable very high throughput. Users of wireless systems also spend a lot on the best available hardware devices in hope of a better systems performance. Unfortunately, the performance obtainable with these carefully designed devices is often degraded by a poor RF planning and ineffective operational Radio Resource Management. Therefore, good Radio Resource Management is required to obtain maximal use of the available spectrum capacity

Not surprisingly, tools for performance optimization are most developed for wireless LANs, the wireless technology widely used as part of the infrastructure of offices, schools, hospitals, and many other business operations and institutions. The means to optimize the performance of such a wireless network are known as Radio Resource Management. Its functions include monitoring the prevailing conditions, collecting this information and changing the operational parameters of the network so as to minimize negative factors and maximize positive factors that affect the overall network performance. Radio Resource Management functions are also useful for wireless network deployment planning.

Although the following is based largely on the problems encountered in the deployment and usage of wireless LANs, much of the discussion applies to other forms of wireless networking as well. Therefore, the generic term "wireless network" is used throughout; the term "wireless LAN" is used if the subject is specific for that technology.

10.1.2 Optimizing Wireless Network Performance

Wireless network performance is affected by two factors: limitations of the available spectrum and, secondly, radio interference due to a variety of causes. License exempt spectrum is a limited resource which has to be shared among the nodes of a wireless network and with neighboring networks. This limitation is a given and is therefore not further considered here. Radio interference arises from the shared use of radio spectrum, largely because spectrum sharing mechanisms like the CSMA/CA protocol are not perfect and because propagation conditions affect their operation. In practice, interference remains an issue because there are many cases where collision avoidance does not work well or even not at all. The notorious hidden terminal problem is a case in point. In addition, there may be other RF devices in the same space that cause interference into the nodes of the network. Typical examples are Bluetooth devices and RFID devices. The significance of Radio Resource Management for a successful deployment of wireless networks can not be over-estimated. This is especially true for scenarios of large or dense deployments such as office buildings, hospitals, and campuses where the co-channel interference is common.[1]

[1] It may be argued that co-channel interference caused by network-own devices is a consequence of limited spectrum being available. Here we take the position that interference that cannot be countered by the network's spectrum sharing features is to be considered as any other interference.

Optimizing the performance of a wireless network requires maximizing the use of the available frequencies and minimizing the interference among the network's nodes, as well minimizing the effects of interference from external sources, i.e. those outside the control of the network operator. Translating these general statements into practical measures requires not only understanding the factors which reduce performance, but also understanding the performance requirements of the users. Radio Resource Management includes establishing the performance goals to be met, as well as the means to determine the actual performance of the network and its components and the means to change the actual performance so as to meet the established goals.

Radio Resource Management usually includes the RF planning and ongoing RF management. RF planning usually covers device placements, initial channel and power assignments for the devices. After the initial planning and network deployment, ongoing monitoring remains necessary. Certain events or conditions may trigger reconfiguration, e.g. the addition of new APs, failures of existing APs, changes in physical environment such as a large metal object being moved into transmission paths, dramatic changes in traffic patterns for the user applications, etc.

10.1.3 Radio Resource Management Goals

The performance goals of Radio Resource Management vary with the purpose of the network and with deployment scenarios.

The most basic objective for a successfully deployed wireless network system is to provide full connectivity in the covered area without any connectivity holes. Given the limited spectrum available, this objective usually requires maximizing the spectral efficiency of the network. In addition, it is also often required to guarantee certain minimum levels of network throughput for the whole system or individual devices. When voice and video applications are involved, there are higher requirements on the wireless network throughput to ensure the desired quality of service (QoS) is achievable. These are the most common goals for Radio Resource Management. The table below summarizes the performance goals that play a role in Radio Resource Management (Table 10.1).

For some wireless networks, maintaining equitable throughput of the access points in the deployment is required, and the Radio Resource Management should take this into account. Note that it is not trivial to determine what is equitable or not. This subject is related to fairness, a consideration usually applied to a heterogeneous population of devices sharing spectrum.[2] In some cases, fairness in channel access probability for each access point is sufficient; while in other cases, fairness can refer to the fair channel occupancy for each device, or fairness in medium time taken for each

[2]See also Chap. 12, Sect. 12.2.4.

Table 10.1 Examples of Radio Resource Management performance goals

	Goal	Comment
1	General connectivity	Assures that wireless coverage is complete without reference to throughput
2	General throughput	Sets a baseline throughput for all network nodes
3	Selective throughput	Sets throughput goals for specific nodes
4	General link reliability	Sets a minimum rate to be maintained at all times and locations
5	Specific link reliability	Sets a minimum rate for specific links to be maintained at all times
6	Load balancing	Assures that the available capacity is shared equally among a set of nodes
7	Stability	Avoids oscillation between different RF plans

device. The bottom line is that an absolute definition of fairness does not exist. What is considered as fair is generally a design choice or a deployment choice.

10.1.4 Inputs: Network Statistics

The factors involved in determining an optimal RF plan include the number of APs, their locations, their PHY types, observed error rates, CCA thresholds, transmitting power levels, the frequency channels, the propagation paths in the deployment environment and the obstructions in the path, external sources of interferences, traffic loads and traffic patterns, variations in clients and client applications, and even the setting of retry limits for each individual AP. Also needed are the statistics of the traffic loads that they generate and the distribution of traffic loads over different levels of Quality of Service. Whether a specific statistic is relevant depends on the Radio Resource Management goals for a given network.

The IEEE 802.11 standard provides facilities to report RF-related statistics in a standard fashion.[3] Because of the dynamic nature of the wireless network environment, such statistics collecting should occur quite often and the RF plans need to adapt to the changes quickly. On the other hand, since wireless operations are nondeterministic in nature, some form of smoothing of the collected statistics is necessary to avoid changes or interruption of the operation of a wireless network too often. A balance between promptness of response and stability of operation is required.

10.1.5 Outputs: Device Configuration Parameters

The Radio Resource Management outputs depend on the capabilities of the network devices. Conventional wireless LAN devices allow setting the operating channel, the

[3] This work was initially performed in Task Group 11k.

RF power output and the defer threshold. They also allow setting of medium access parameters such as the baseline contention window size and the AIFS for the different QoS service levels. With these parameters, device behavior can be influenced such that generally the network's performance objectives are met or adequately approximated.

10.1.6 Summary

The preceding sections provide an overview of Radio Resource Management, its goals and its operations. In the following sections, we will first discuss in detail the challenges in generating an optimal RF plan in a timely manner. Then, we will describe the major categories of Radio Resource Management and some of the major differences in functionality. The chapter ends with a discussion of two major approaches for Radio Resource Management: the distributed Radio Resource Management approach and the centralized Radio Resource Management approach.

10.2 Common Issues of Radio Resource Management

10.2.1 Complexity

Radio Resource Management is notoriously difficult due to its high degree of complexity, its dynamic nature, variations in deployment scenarios, and operating conditions, etc. The degree of complexity is made clear by the following example: for a wireless network consisting of N_{ap} access points, N_{ch} available non-overlapping channels and N_{pwr} power levels per access point, there are:

$$\left(N_{ch} \times N_{pwr}\right)^{Nap} \tag{10.1}$$

possible permutations of RF plans. For $N_{ap} = 20$, $N_{ch} = 3$, and $N_{pwr} = 4$, for example, the total number of RF plans is extremely large, about 4×10^{21}. This suggests that the problem of optimum channel assignment is in fact an NP-complete problem[4] that requires approximations to obtain a solution.

Even if the results of interest are limited to RF plans that optimize a single variable, say, the overall network throughput, the number of possible outcomes is still very large. Assume that we only consider two controllable RF parameters: the channel assignment and the power level. Clearly, this leaves out many factors that contribute to the overall system throughput. Let T(P) be the overall throughput for the whole wireless network corresponding to an RF plan represented by P. We have a

[4] See Garey and Johnson [131].

wireless network consisting of N_{ap} access points with the same type of radio, and N_{ch} available non-overlapping channels. The Radio Resource Management goal is to find the optimal RF plan of $\{(\text{channel, power})_s : s \in S\}$ that corresponds to the maximum system throughput T.

$$T(P) = \sum_{s=0}^{Nap} \frac{S}{n(s)} \tag{10.2}$$

where n(s) is the number of co-channel APs in interference range for the access point *s* and *S* is the single access point throughput when it takes the entire channel without any interference.

If one searches the optimal RF plan by enumerating all the possible plans, it can take days of computing time for a wireless network consisting of, for example, 100 APs. Because of such a high degree of complexity, it is beyond human capabilities to manually tune the RF parameters, unless the deployment is a very small. Therefore, automation is required.

Much research has been done in this area, but notable success has proved elusive. The trend towards dynamic spectrum access further complicates this issue. In some frequency bands, sharing with primary spectrum users may require the implementation of dynamic spectrum sharing procedures, involving monitoring as well as changes in operating channel. The most obvious example is DFS, the dynamic frequency selection,[5] required in the 5 GHz bands, but operation in TV White Space bands may have similar consequences.

10.2.2 Responsiveness

Many events can occur in a wireless network that affect the throughput or another performance criterion so much that the current RF plan must be re-evaluated and, if necessary, changed. Examples of such events are:

- an access point that becomes not functioning, which leaves a hole in the coverage
- the addition of an access point to extend coverage
- changes in the characteristics of RF propagation path
- a moving object can stand in a RF propagation path
- a burst of external interference triggering a DFS channel change

All these events can happen at any time and Radio Resource Management should respond rapidly to these dynamic events to keep the wireless network optimal. Optimizing an RF plan for a particular scenario can take a long time. Therefore, updating the operational RF plan may lag behind the external causes. As a consequence, the wireless network may have to operate sub-optimally for some time.

[5] See Chap. 9, Sect. 9.1 for more detail on DFS operations.

10.2.3 Automation and Convergence Speed

Because it is beyond the capability of a network administrator to deal with the high degree of complexity of the Radio Resource Management decision making, automation is necessary.

Automation of Radio Resource Management is still in its infancy. Commercial offerings of automated Radio Resource Management may show a performance that is far from being satisfactory and that leaves much room for improvement.

As shown in the example in Sect. 10.2.1, RF plan generation can be treated as a classic search for the optimum of a target function for a very large number of configurations. The complexity of the searching process is exponential with respect to the number of access points. A brute search can take several days to reach the optimum of the target for a wireless network with hundreds of access points. With so many factors involved, it is important to decide which factors are the most important. Ideally, the optimization procedure should capture the salient nature of the deployment's issues with the least number of factors, but without sacrificing the accuracy of the resulting RF plans. On the other hand, less significant factors in one deployment can become important if the deployment changes. It therefore requires a lot of discretionary decisions to determine the most important factors for a particular deployment.

There have been many research projects on the application of optimization algorithms, for Radio Resource Management, especially the choice of a channel assignment scheme. Graphical Edge Coloring or Map Coloring, among them, are probably the most researched methods for deriving the optimal channel assignment plan. There have been many attempts in practice to approximate and speed up the automated search processes. For example, one may choose to cut off the search process after running some pre-determined length of time and simply use the local optimum found so far as the resulting RF plan. However, there is reason to assume such a RF plan to be optimal. Other attempts roughly identify the likely region of an optimum with some empirical approaches followed with a fine searching within that region. While such an approach could greatly reduce the searching time, it could also rely too much on human experience for the empirical selection of the likely best region. Therefore, it can easily miss the real optimum or be stuck with a local optimum.

10.2.4 Flexibility of Radio Resource Management Goals

Depending on the users or the applications of a wireless network, network performance goals vary and may include different factors, e.g. as given in Table 10.1. Some organizations or applications may require that performance goals can be set so as to assign more bandwidth to some particular set of devices or some application. The Radio Resource Management target function should be flexible enough to

accommodate such special needs which may be built into the target function in some analytic form.

10.2.5 Selection of the Radio Resource Management Target Function

The choice of a particular analytic form of the target function is quite arbitrary. The non-uniqueness leads to further questions of matching an optimal RF plan to the optimal wireless network performance. Further, the weight or the form of a particular factor expressed in the target function affects the resulting RF plan directly. Therefore, whether an RF plan is optimal or not can become quite subjective, depending a lot on how a target function is determined.

For example, in the target function (Eq. 10.1) in our example, we treat all the neighbors of an access point the same. However, in practice, not all neighbors are necessarily equal. A neighbor with very light traffic has less impact than those with heavy traffic. If we take this into account and add a weight *w(s)* to each neighboring access points, we can come up with a new form for the target function:

$$T(P) = \sum_{s=0}^{Nap} \frac{S}{n(s)} \mathrm{w(s)} \tag{10.3}$$

The empirical nature of the choice of Radio Resource Management target function leaves a lot of room for network management expertise, as well as for innovation and improvement.

For a network carrying light traffic, a sub-optimal RF plan may not cause a major problem. This is typically the case in many network deployments, whether domestic or institutional. However, with bandwidth consuming applications such as video becoming more and more popular, operational throughput demands increase, and this makes a sub-optimally configured network unacceptable. Therefore, Radio Resource Management that leads to optimal wireless network performance in the real world becomes more desirable.

10.2.6 Selection of Relevant of Performance Metrics

Good performance metrics are essential ingredients for a good RF plan. For the same form of the target function, the resulting optimal RF plan can still vary with the values chosen for the many thresholds involved. For example, if one chooses a different threshold value to determine whether an access point is a neighbor, the resulting RF plan will be different. In addition, many factors contributing to the wireless network performance have statistical distributions, such as traffic loads, traffic patterns, frame error rates, medium access parameters, and frame retries.

These statistical factors add complexity and must be applied carefully so as to maximize the quality of the RF plan.

Matching the Radio Resource Management target function with the real world performance is non-trivial. As so many factors contribute to the resulting network performance, it is very difficult, if not impossible, to come up with an analytical or empirical form of Radio Resource Management target function that corresponds to the real world network performance. A remedy for this fundamental issue is urgently needed. One emerging and promising approach is to use real-time simulations, including protocol level medium access simulations, to determine the performance metrics to be used in generating a given RF plan. When using the real-time statistics from the real world wireless network as inputs to the simulations, the predictions can be accurate and have a clear correlation to the real-world performance.

10.2.7 Radio Resource Management Information Exchange

In general, Radio Resource Management requires the exchange of channel information and for passing management information. This exchange can make use of a dedicated communications channel or radio. A multiple radio capability can lead to simple and efficient Radio Resource Management operations. On the other hand, a dedicated radio increases the hardware cost and power consumption, which is a non-trivial burden. Some Radio Resource Management research suggests adopting a dedicated common control channel (CCC) to transmit the control and management frames for RF management information. If such a common control channel makes use of a dedicated radio, it helps to improve the efficiency in Radio Resource Management and increases the MAC efficiency for the channels of data transmission. The risk is that the common control channel can become a bottle neck of communications if there is much interference or, worse, some malicious jamming, in that channel. Further, the control channel is not guaranteed to be available and this leads to the possibility of a single point of failure. The impact of failure on the overall wireless network could be too devastating to afford. This may be the top most concern that prevents a common control channel for Radio Resource Management from being adopted. In addition, taking away a channel from data communications can also have non-ignorable impact on the overall system performance. For 2.4 GHz for example, when the available channels are already limited to three non-overlapping channels, such an overhead is hardly affordable.

10.2.8 Forms of Radio Resource Management

Radio Resource Management can be implemented either in centralized or distributed form. Centralized Radio Resource Management generates RF plans and other configurations, based on the data collected from the network nodes, and distributes

the results to these nodes. The location of such a Radio Resource Management facility is not immaterial: it must support the communications capacity necessary to collect Radio Resource Management information and distribute Radio Resource Management control commands.

Distributed Radio Resource Management, on the other hand, does not require a central entity. Instead, each wireless device develops its own "RF Plan" based on its own observations and, if available, the information obtained from nearby devices. However, because each device' view of the network is limited to its own neighborhood, the resulting RF plans may not add up to a (near) optimum overall RF Plan.

10.2.9 Summary

Radio Resource Management is core functionality for any major wireless network management system, yet it is the least mature. Some fundamental issues, such as the NP-completeness of e.g. the problem of channel assignment, the correspondence between a Radio Resource Management target function and network performance in the real world, remain unsolved. Until these fundamental issues are resolved, Radio Resource Management will continue to be among the most promising areas of wireless technology research.

In the next sections, we will look into these Radio Resource Management approaches and discuss their advantages and downsides. The discussion will be focused on indoor wireless networks, but much applies to other scenarios such as wireless Mesh networks or dense Femtocell deployments.

10.3 Centralized Radio Resource Management

10.3.1 Introduction

Centralized Radio Resource Management monitors and controls the wireless parts of the network from a central entity, which can be either a standalone network controller or a function of a network controller. Centralized Radio Resource Management often uses wired connections to collect status and statistics from wireless devices under its management and to send them control commands. Where such links are not available – e.g. in mesh-like deployment, wireless connections have to be used. There are quite a few RF management systems available on the market taking the centralized approach: most enterprise wireless LANs adopt this approach, regardless of their approach to data switching and routing.

Centralized Radio Resource Management has a number of advantages. It tends to be more efficient than a distributed Radio Resource Management because no multiple rounds of iterations within the entire wireless networks are required to obtain a stable RF plan. Since the local network and channel conditions for each device are often transmitted over the wired path, there is little overhead of management

frames for Radio Resource Management on the wireless medium. This leaves more bandwidth to the wireless data traffic. The Radio Resource Management decision made by a centralized Radio Resource Management is sometimes better and obtained faster than that of distributed Radio Resource Management. The reason is mostly that the former has a global view of the entire wireless network and thus more complete information than the latter. Further, there are fewer requirements on the wireless devices in case of centralized RF management and, therefore, the implementation cost of these devices can be lower.

On the other hand, centralized Radio Resource Management has some major drawbacks. Because centralized Radio Resource Management has the inherent difficulty of NP-completeness in complexity, the number of devices each centralized Radio Resource Management systems can handle is bounded and, therefore, scalability is a major concern for a centralized Radio Resource Management. Nonetheless, thanks to various approximation algorithms that have been developed, centralized Radio Resource Management systems are able to generate acceptable RF plans in a timely manner. Another disadvantage is the controller or server – which has to be duplicated to assure continuity – is a significant extra cost factor, especially for small deployments.

10.3.2 Algorithms

As discussed previously, in a good approximation-based Radio Resource Management, channel assignment has to be both fast and close to the optimum. Speed can be achieved at the expense of optimality: when speed results from reduced run times, for example, the quality of the resulting RF plan quality is likely to be questionable and poor wireless network performance may be the result.

A good Radio Resource Management algorithm is not easy to design; even less so if it has to be fast. These algorithms promise to be a significant area of innovation in the near future. The following looks at the state of the art.

Channel assignment is at the core of the Radio Resource Management and it is often used as example of the degree of complexity for Radio Resource Management problems and algorithms. Many research projects have tried to apply Graphical Coloring or Map Coloring to channel assignment.

In Graph Theory, a k-coloring of an undirected graph G = (V, E) is a function

$$c : V \rightarrow (1, 2, ..k) \tag{10.4}$$

such that c(u) does not equal to c(v) for every edge (u, v) belonging to E, where the numbers (1, 2,…k) represent k colors, E and V the edges and vertices in the graph respectively.

Adjacent vertices must be colored with different colors.[6] The problem is to find the minimal k, the minimum number of colors for a k-coloring. The vertex-coloring

[6]Jensen and Toft [71].

problem can be proven to be NP-complete in complexity, a type of problems widely suspected to have no polynomial-time algorithms for solving them.

When applied to the problem of channel assignment, wireless devices are represented by vertices in a graph. Two devices are connected with an edge in a graph if they are close enough to be within transmission or interference range of each other, based on any pre-determined signal strength threshold. The available channels for a device are represented by colors. The graph-coloring problem is to find the minimum number of colors satisfying the coloring constraint that any pair of connected vertices (= all edges) must be colored differently. However, wireless devices, in fact, have only limited number of channels, or colors. As a result, the coloring constraint in general cannot be satisfied. So a channel assignment graph-coloring problem, in fact, does not directly correspond to a classic graph-coloring problem. Rather, it corresponds to a variation of graph-coloring problem, one with various optimization goals, such as maximizing the overall throughput or maximizing link throughput. In general, graph-coloring with optimization goals is NP-hard[7] in complexity. Although it is unlikely to find a polynomial-time algorithm to solve the channel assignment problem exactly, it is possible to find approaches to work around the NP-completeness property. When the size of deployments is not too large, an algorithm to deal with exponentially increasing run time may be acceptable. Even for a large deployment, one can still find near-optimal approximation approaches that run in polynomial time, which are often good enough in practice.

A commonly cited and researched approximation approach is to apply a Greedy Algorithm[8] to channel assignment. A Greedy Algorithm divides the decision making process to find the optimum into a series of intermediate steps referred as local decisions. It makes a locally optimal decision in the hope that a series of locally optimal choices will lead to a globally optimal solution. It runs very fast. But how its optimum relates to an optimal wireless network performance metric remains unproven. Therefore, a Greedy Algorithm is pretty much an empirical approach. The quality of the RF plans generated with a Greedy Algorithm indeed depends heavily on the experience of the designers and their treatments of various factors for the performance target function.

When applied to the channel assignment in Radio Resource Management, a Greedy Algorithm generates a RF plan with a series of "greedy" decisions, which maximize the Radio Resource Management metric at each step. It assigns channels to devices one at a time, until all the devices are assigned. It takes two steps for each assignment: the *device selection* step to decide which device from the set of the unassigned devices to assign a channel to next, and the *channel selection* step to decide what channel to assign. Decisions are made at both steps to optimize their respective predetermined matrices.

[7] See Garey and Johnson [131].

[8] See Cormen et al. [132].

Below is an example of the Greedy Algorithm for Radio Resource Management, which assumes that:

- C = { c1, c2… }: the set of candidate channels for the devices;
- S = { s1, s2… }: the set of all the devices;
- U: the set of unassigned devices;
- A: the set of assigned devices;
- α(A): the pre-determined function deciding the next device to assign;
- β (A, CH): the Radio Resource Management target function devices given a set of assigned devices {A} and their assigned channels {CH};
- CH, the RF plan for assigned devices, consisting of pairs of a device and its assigned channel,

The algorithm consists of the following steps:

```
A ← { }, CH ← { }, U ← S
while U ≠ { } do
    Select s* such that α(s*) = max (α (s) : s ∈ U)
    Select c* such that β (A∪ {s*}, CH ∪ {c*}) = max (β (A∪ {s*}, c) , CH ∪ {c}: c ∈ C)
    U ← U – {s*} , A ← A ∪ {s*}
    CH ← CH ∪ {(s*, c*)}
end while
return CH
```

In the end, A = S, U = { }, and β (A, CH) where β corresponds to the optimal performance metric and CH the resulting channel assignment.

It is noted that the set of C could vary for each device, but are made the same for all the devices to simplify the explanation. We only include the channel assignment as the example of the Greedy Algorithm. It can indeed include other factors such as power level assignment, etc. The algorithm is very fast. Its complexity is $O(N^2)$, where N is the number of devices. Greedy Algorithms often lead to optimal solutions, but in general there is no guarantee to tell. There are many ways to choose α and β. For example, α can be as simple as the number of neighbors for the devices with assigned channels:

$$\alpha(s) = \sum\nolimits_{s' \in A} I(s, s') \tag{10.5}$$

where I (s, s') = 1 when s and s' are neighbors and is 0 otherwise.

$$\beta = \sum\nolimits_{s' \in A} S(s) \tag{10.6}$$

where S is the throughput for the device s.

The quality of the RF plan generated with a Greedy Algorithm relies on how α and β are chosen. It demands a lot of experience and careful design considerations to come up with a good choice of a particular performance target for a real wireless network. Finally it is noted that Radio Resource Management operations that change the RF channel plan may require significant hysteresis to avoid oscillations. The latter could

result from changes in traffic patterns or propagation conditions. Avoiding oscillations facilitates the emergence of a stable "inter-work sharing pattern" at network boundaries. The same requirement applies to Distributed Radio Resource Management.

10.4 Distributed Radio Resource Management

10.4.1 Introduction

Distributed Radio Resource Management does not require a central entity to collect network operation information and to distribute network control information. Instead, each wireless network device chooses its channel, power levels and other configuration parameters on its own, based on the local conditions it observes, and, if available, it will use related information received from other devices.

Distributed Radio Resource Management has a number of important advantages. Most significant are its inherent scalability and robustness. A distributed Radio Resource Management provides an effective approximation to work around the NP-completeness of complexity in channel assignment. This approximation is inherent in the limited extent of its domain of operation, even in the case of cooperative Radio Resource Management. Another advantage is the robustness: if one network node fails for some reason, this does not affect the nodes around it, except for a possible increase in traffic load to accommodate the capacity loss of the failed neighbor node. A wireless network with distributed Radio Resource Management is also not as costly as a centralized Radio Resource Management, since a centralized Radio Resource Management server or function is not required; this is an advantage that many small or medium deployment owners appreciate.

Distributed Radio Resource Management can be either non-cooperative or cooperative. In either case, the operation is continuous in the sense that the local Radio Resource Management entity must monitor its domain and take action if changes affect its performance objectives.

In the non-cooperative case, each node in the network determines its own optimum channel parameters, based on its own observations of channel conditions and operational requirements such as offered load.

In the cooperative case, a device will actively share with other devices its operating parameters and local statistics, as well as other information it may collect. A cooperative approach facilitates a more comprehensive view of the variables affecting local performance and, therefore, it may be able to yield a better local RF plan than non-cooperative Radio Resource Management, which is based solely on its own local observations. However, the cooperative approach introduces overhead for a device to sense or probe the channel and pass around the information to other devices. This not only imposes an additional complexity in implementation, but also consumes medium bandwidth and lowers channel utilization efficiency. Several iterations may be needed for the entire wireless network to reach a stable

channel assignment or an RF plan. A distributed Radio Resource Management, therefore, often takes some time to reach stability and it may be slow in response to the changes in channel conditions and network conditions. When a node has selected its optimum channel, the selection will affect the choices of the devices around it. Their choices, in turn, will feedback into the device's next round of sensing and adjustment, etc. Therefore, distributed Radio Resource Management is a wavelike process that may take a number of iterations before a stable RF plan is has been implemented.

This can become a concern in case changes in wireless channel conditions or network conditions frequently trigger re-evaluation and regeneration of the RF plan. An additional drawback is that the frames for information exchanges can be lost, causing the information to be used for a decision incomplete or out of time. Finally, it should be also noted that unfairness could occur to non-supporting devices, which is true of all Radio Resource Management approaches.

10.4.2 Design Considerations

Design considerations for cooperative distributed Radio Resource Management include:

- Definition of the form of the target function and the factors to be built into it
- Data types and formats for Radio Resource Management input statistics and output commands
- Choice of standardized (e.g. IEEE 802.11 "k") or proprietary frame formats
- Definition of the protocol for transmitting Radio Resource Management information
- Definition of the timeout thresholds for statistics collecting
- Definition of the State Diagram for the Radio Resource Management engine

Figure 10.1 shows a much simplified example of a State Diagram for a distributed Radio Resource Management entity. It scans the channel in the "Scan State," decides the best channel (usually the one with the "least cost" measured against some criterion, e.g. transmit time) and goes to the "Run State" once the channel is selected and activated. In the Run State, performance is monitored and if it changes, the Scan state may be triggered to re-evaluate the RF plan to decide the best channel under the new conditions. This process is applied separately to all links of the node that do not share the same RF channel.

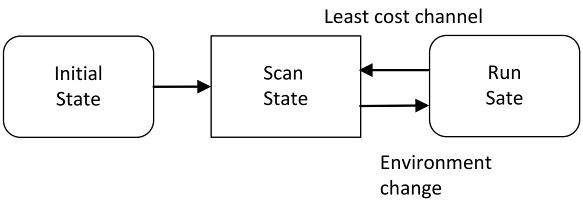

Fig. 10.1 A state diagram for distributed radio resource management

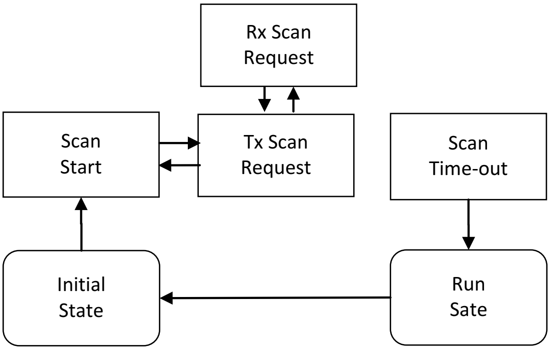

Fig. 10.2 Event sequence diagram

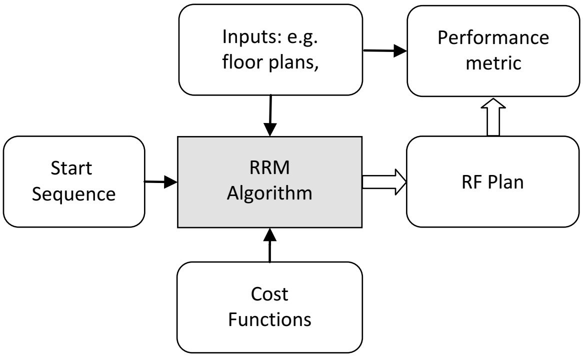

Fig. 10.3 System flow diagram for a distributed radio resource management

Usually, the RF plan generated by a distributed Radio Resource Management is not as optimal as a plan generated with a centralized Radio Resource Management. This is due to lack of complete information for the entire wireless network. The main concern is how close the resulting RF plan is to the optimum of the wireless network performance in the real world. However, this is also a concern for any other approximating approaches that work around the NP-completeness of the channel assignment problem.

The following example uses the target function in Eq. 10.2 to formulate a simple distributed cooperative Radio Resource Management algorithm for a network consisting of 16 nodes. The number of all possible RF plans is exponential to the number of devices, 16 in this case. The algorithm in this example is simple, as illustrated in Fig. 10.2; it runs on each device individually. It involves scanning the channel for local observations and receiving information from other devices. The device also sends out the information it collects to other devices. When the scanning period is over, the device uses information, both local and transmitted from other devices, to calculate the "cost" for each channel, based on the pre-determined Radio Resource Management target function, and selects the channel corresponding to the least cost. This process is illustrated in Fig. 10.3.

To assess the quality of a generated RF plan, it is necessary to enumerate all possible RF plans and select from them the one corresponding to the maximum overall

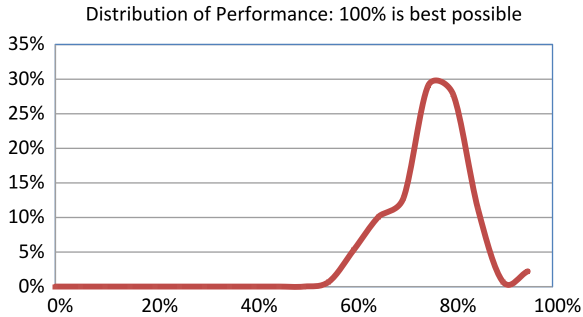

Fig. 10.4 Percentage distribution of relative performance

system throughput. Dividing the throughput for each individual RF plan with the maximum performance metric gives a normalized performance metric, called the relative performance.

It will be clear that the specifics of the chosen RF plan depend on the sequence in which individual devices run their Radio Resource Management operation. To what extent the resulting RF plan depends on the starting sequence is an important indicator of the quality of the distributed Radio Resource Management algorithm. One way to examine this dependency is to simulate the distributed Radio Resource Management process with a large number of permutations of starting sequences and show the relative performance distribution. A good algorithm should only have a weak dependency, whose distribution should be a narrow and sharp peak corresponding to a high relative performance.

Figure 10.4 shows the relative performance distribution of the simple distributed channel assignment algorithm. It shows the percentage of the Radio Resource Management plans versus their relative performance. These Radio Resource Management plans are sampled based on various starting sequence for the devices to start their Radio Resource Management. One can observe that most resulting RF plans give a relative percentage of performance at around 80%, indicating that such an algorithm has a weak dependence on the order of starting sequence and the resulting RF plan is close to be optimal.

The question remains how relevant the target function is with respect to the performance of the network in the real world. Many factors should be taken into account for the Radio Resource Management target function. In our example, the target function is overly simplified and it should not be surprising that the optimum so generated does not really lead to an optimal performance in the real world. Even with a more sophisticated and comprehensive target function that has all the factors one can think of built into it, the correspondence between the generated optimal RF plan and the optimal performance in the real network will still be in question, for many of the reasons discussed in Sect. 10.2.

One likely remedy, and perhaps the only one, that can work around the inherent difficulties in formulating a target function corresponding to the real world performance is to run a full blown wireless network simulation to generate the performance metric, the overall system throughput, for example, using the real-world

statistics as inputs, rather than assuming an unrealistic analytical metric, such as MAC level efficiency, based on some overly simplified assumptions. When the inputs used are real-world statistics, such a simulation-based approach can be very powerful in accurately predicting optimum performance goals.

10.5 Cognitive Radio and Radio Resource Management

The preceding sections address Radio Resource Management in its current, conventional setting: a wireless network that operates in a given frequency band, possibly adjacent to other networks. Here the world is not necessarily completely static, but changes are limited to interaction between network elements at the network boundaries. Achieving stability is relatively easy.

Cognitive Radio technology introduces a new and different element in this picture by removing the "static" channel arrangement and replacing it with a temporary one that will change with time. Since the core concept of Cognitive Radio is to use spectrum that is not in use, a reduction in a network's traffic pattern could lead to loss of spectrum to a nearby system. If this is the case, it is in the interests of network operators to prevent such loss – if necessary through "keep alive" traffic that serves no other purpose than securing spectrum access. On the other hand, keeping the rate of resource changes low enough to assure stable spectrum resources may conflict with the reason for deploying Cognitive Radio in a given network.

This short analysis suggests that Cognitive Radio and optimal spectrum efficiency through Radio Resource Management do not necessarily go together.

10.6 Conclusions

As we have discussed, the quality of Radio Resource Management is critical to the operation of wireless network deployment, especially for enterprise class wireless LAN systems. Issues arising from poor radio resource planning or lack of wireless network optimization are among the top customer complaints for large wireless LAN systems. The ever growing density of wireless LAN deployments, as well as more wireless network traffic generated by end users adopting bandwidth consuming applications, like videos, exacerbates sub-optimal configuration choices and ditto channel plans.

It is very challenging to provide good solutions that take into account many nontrivial factors such as interference, dynamic channel conditions, traffic loads, class of services, and user prioritization. Therefore, this area offers exciting opportunities for innovations that will deliver significant performance improvements in wireless networks.

Part III
Prospects

Following the preceding forays into the theoretical background of spectrum sharing and into its complex practice, the concluding chapters of this book look at the future – at the emerging applications of wireless technologies as well as at the possible evolution of spectrum sharing research and practice.

Chapter 11
Emerging Applications of Wireless Technologies

Smart systems are beginning to pervade society at many levels, which go well beyond the ubiquitous smart phone that knows where it is and what its owner is doing. One example is the use of movement sensitive smart thermostats and light switches that reduce the cost of climate control and lighting; another example is the communication systems in cars that alert emergency services in case of accidents. The list is potentially endless. The common element in this diverse picture is that the smarts must be complemented by connectivity to function effectively and yield maximum benefits. Since many smart systems are non-stationary, wireless connectivity is necessary. The following addresses three areas of application of wireless technology as examples: the internet of things, the smart grid and intelligent transportation systems.

11.1 The Internet of Things

The "internet of things," (IoT) as a phrase, dates from the last years of the twentieth century. It denotes the concept of a pervasive intelligence and communications ranging from simple sensors to large server-based networks that cooperatively control whole industrial facilities. The number of devices involved may run into the 100s of billions. The IoT concept and its potential for economic benefits, as well as technical and legal challenges, have fired the imagination of many, including the European Commission.[1] It may be argued that IoT is already being implemented in various forms in different industries. With this in mind, the IoT concept is likely to remain what it is today, namely a convenient catch phrase for a wide variety of devices, services, and applications. Although conceived as a vision, its growth will be driven by economic factors, which differ by sector and industry.

[1] See e.g. EC 2009, *Internet of Things – An action plan for Europe* – it provides interesting insights as well as many references.

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_11, © Springer Science+Business Media B.V. 2011

Relevant to the subject of this book is the long-term prospect of interconnection at scales ranging from a few bits per second over a few meters distance to high speed links able to deliver real-time video and control to and from mobile objects. The need to assure spectrum availability has been recognized by the FCC, as well as the EC and other regulatory authorities. But the demand for new spectrum to support the growth of IoT applications is hard to gauge, largely due to the variety of data rates and transmission distances.

An important element in the IoT landscape is Machine to Machine communications (M2M). Standards are being developed with the goal to provide a unified architecture for the interconnection of sensors, actors, and server systems that are used in IoT applications. Many cellular service providers see M2M as a major new market opportunity. The M2M architecture[2] being developed shows that many M2M deployments are expected to make use of short-range, possibly proprietary, radio links – which are sometimes referred to as capillary links – whereas cellular networks are expected to be the means of choice for wide area communications. Assuming this expectation is correct, the capillary links will be the area of IoT technology in which spectrum sharing will play a key role.

11.2 The Smart Grid

"Smart Grid" is another example of application using wireless technology that has the potential of subtle impacts on people's everyday life. The term "Smart Grid" refers to electrical utility systems and networks enhanced with broad collection of functions and capabilities, which improve generation, distribution, and delivery of electricity. The current electricity grids are highly centralized structures designed around central generation and distributed consumption of energy. As society becomes more and more energy intensive, variations in energy usage have increased to a point where generation and distribution systems have difficulties in assuring a constant level of service. The emergence of intermittent, decentralized generation further complicates matters. Today's utility networks are only partially automated[3]; much is expected of making these grids "smart," which means being able to adapt to varying supply and load on every scale, from the home equipped with solar arrays to the nationwide utility network which distributes electricity to its clients. In terms of the IoT concept, smart grid is a collection of IoT applications operating at different levels and with different purposes. A common element is the M2M communications model, which defines three levels:

- Capillary networks at the energy consumer location collect and distribute energy consumption and control information.

[2] See ETSI portal under M2M.

[3] This varies by region and country. Notably Europe is well advanced in automation of its large scale electricity transport networks, unlike many local distribution networks.

- Wide area networks connect consumers to the energy distribution networks.
- Utility core networks which connect large-scale generation and transport facilities.

In the capillary networks, spectrum sharing among dissimilar technologies will play a key role.

11.3 Intelligent Transportation Systems

A parallel development can be seen in the world of transportation – it is known as "Intelligent Transportation Systems" or ITS.[4] ITS promises improved traffic management and efficiency, as well as improved transportation safety through driver information and possibly driver assistance. Given the mobility of vehicles, wireless communications are a necessity for vehicle to roadside infrastructure ("V2I") and for vehicle-to-vehicle (V2V) connectivity. The former focuses on information exchange between the road infrastructure and the vehicle. The applications vary from road toll collection to emergency signaling and traffic flow control in various forms. Various forms of communication are being considered for V2I, including short-range communications called WAVE,[5] which are based on IEEE 802.11 technology. The WAVE communications protocol is application specific which includes forwarding capabilities. It relies on a simplified form of the IEEE 802.11 CSMA/CA medium access protocol where only the collision avoidance feature is used, but neither the exponential back-off nor the Acknowledgement. Vehicle-to-vehicle communications pose many challenges and a fertile area of research.

11.4 Femtocell Ubiquity

As cellular networks expand in scope and in services offered, network coverage and capacity have to increase. The femtocell represents an emerging low-cost application providing improvements to both coverage and capacity. Femtocells are small in-building cellular base stations, which are user installed and that connect back to the operator's network through an ADSL, cable, or fibre connection. The latter becomes the limiting factor in the capacity seen by users, not the link budget of the outdoor macrocell network. Because of the RF isolation provided by building walls, these femtocells can re-use the spectrum used by the macro cells of the outdoor network. The cost of a femtocell-based network "build-out" is a fraction of the cost

[4] In the US, the Research and Innovation Technology Administration is a key player in this field. In Europe, ETSI is developing standards for ITS. There are a number of car industry consortia active in this field as well.

[5] WAVE = Wireless Access in Vehicular Environments.

of a macrocell build out that achieves the same capacity increase. However, if femtocells really become ubiquitous and penetrate dense urban complexes, the building isolation may prove inadequate, and interference management will become an issue that puts spectrum sharing in these bands on the front pages of the trade press.

11.5 Conclusion

There is no reason to doubt that wireless technologies of all kinds will continue to grow in importance and in the breadth and depth of their use, domestically, institutionally and publicly. Given the physics involved – notably the Shannon-Hartley theorem and the variability of the environment – which constrain spectrum use, spectrum sharing is assured of a rich and varied future. Technological developments will continue to chafe at these constraints, and pundits will keep saying that they are irrelevant. Ultimately, practice will reveal which clothes the various players wear.

Chapter 12
The Future of RF Spectrum Sharing

RF spectrum sharing is a subject with a long history and an even longer future. Its prospects vary from relaxing the current limitations on cellular capacity to freeing spectrum for use by all, at anytime, anywhere. Free spectrum access purportedly will lead to greater prosperity for society as a whole. On the other hand, practice suggested that there is much to learn for all concerned before that future potential can be realized. Naysayers point out that it is not even clear that any of the claimed potential can be realized at all. As with most subjects, predictions, whether negative or positive, tend to overshoot; therefore, it is necessary to look at the boundaries that will shape future developments. Using three key wireless technologies as examples, this concluding chapter first looks at the current status of wireless systems and spectrum usage. It then discusses the expected growth of demand for wireless capacity. It finally looks at possible ways to address that growing demand, notably in ways that include spectrum sharing solutions.

12.1 Where We Are

12.1.1 The User Perspective

Users are largely ignorant of matters related to RF spectrum. Even the terms themselves many find hard to relate to and the subject is difficult to comprehend. On the other hand, most users of cell phones and wireless network have experienced the effects of spectrum overloading at one time or another without being aware of the causes. Large sports events and major calamities frequently cause a temporary or intermittent system overload because of the scarcity of available bandwidth, as did the release of the first iPhone, even in the licensed spectrum. Many new iPhone owners wanted to try out emails and web access on their new

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1_12,

phones and the AT&T network was not able to handle the amount of air traffic generated by these phones.[1]

Another more tangible instance of user perceived impact of spectrum sharing is the deployment of wireless LAN networks in major cities. San Francisco is a well-known example. During weekdays and nights, the wireless LAN performance is generally adequate. But in the morning and evening hours, service tends to be sluggish at best. The extra capacity demand of morning and evening hours exceeds the capacity of the 2.4 GHz band which offers only three non-overlapping channels. As wireless LAN signals easily spread throughout a group of multi-story buildings, one can easily observe more than ten wireless LAN access points at any point inside or outside a house or apartment. Interference due to asymmetrical downlinks and uplinks leads to the access points reducing their data rate and thus increasing the time they transmit, which in turn results in a higher collision probability. The channel capacity is thus quickly exhausted. Similar effects are observed at major sports events and at large conferences.

Although people are getting used to the quirks of "wirelessness," they prefer reliable operation of their gear, regardless of the services used and density of active devices. It is up to the wireless research community to address the underlying issues and to help the wireless industry meet the demands of its users.

12.1.2 The Regulatory Perspective

Spectrum managers and regulators are clearly aware of the potential benefits of liberalizing spectrum usage. They have carefully implemented liberalization policies and they have avoided the pitfalls of opening up too much spectrum to too many users with little or no understanding of the potential consequences. These policies have come under attack, mostly from pundits and special interests, as being too careful and too slowly implemented. The proponents of spectrum liberalization are motivated by the belief that the market is better in allocating scarce resources than the regulator, irrespective of the latter being a commercial spectrum manager or a government authority. Practice in other industries suggests that this simple dictum is a little more than that: a statement rather than an immutable law. Therefore, the restraint with regard to spectrum liberalization shown by regulatory authorities is commendable.

The key policy issue for regulatory authorities is to avoid the creation of a practice that cannot be rolled back in case severe problems arise. This is notably true in case of license exempt spectrum: once a regime with conditions and requirements for its use have been defined and industry has developed commodity products based on that regime, a broadly acceptable way back may be hard to find.

[1]Many new iPhone owners wanted to try out web access on their new phones and the AT&T network was not able to handle the amount of air traffic generated by these phones. Corrections in the iPhone software reduced the burden on the network and removed the problem.

In the case of tradable licensed spectrum, which is sub-licensed to the actual users, the number of participants is low and the spectrum manager can determine the conditions and requirements for spectrum access. Further, the sub-licenses will be time limited and they will carry sanctions in case the sub-licensee does not comply with the rules set by the spectrum owner. Thus, not only are potential problems less likely, but corrective action, if necessary, is effective and easily implemented. The same applies to the exclusively licensed spectrum. Even under the general framework of service neutrality and technology neutrality, setting rules for spectrum users that allow different technologies to be operated in the same frequency band is already common practice.[2]

These considerations leave the license exempt spectrum as a potential tinderbox whose fire can bring the wealth of Wi-Fi or the stormy protests of the tragedy of the commons. In any case, it will be entrepreneurs who develop new products, who will clamour for more cheap and open spectrum. Keeping them out of each other's harming ways requires medium access criteria that are less myopic than listen-before-talk. This chapter ends with a suggestion towards such criteria (in Sect. 12.3).

12.1.3 The Industry Perspective

The wireless industry has developed and implemented many schemes of spectrum sharing, ranging from the tight in-band and out-of-band emission limits and block edge masks of the cellular industry to the carrier sense multiple access scheme of wireless LANs. All of these schemes focus on optimizing the system itself rather than aiming at generic spectrum sharing. Such a policy makes sense, given that the current radio regulations are selective rather than inclusive.

Even in the case of inclusive spectrum regulations such as those applying to the ISM bands, the wireless industry has been loath to implement generic spectrum sharing mechanisms because spectrum sharing is intuitively seen as a zero sum game where sharing spectrum means losing something: e.g. operating range, throughput, or quality of service. Although this is correct in principle, in practice this loss may not be noticeable because actual traffic load does not always approach capacity. Therefore, users may well prefer systems that share spectrum efficiently and invisibly.

The next problem is the lack of a broad theory of spectrum sharing that allows those involved to make informed decisions. Instead, debate about the why and how of spectrum sharing between dissimilar systems tends to circle around the perceived need to defend the own system or design from ill-defined threats. The lack of progress is, therefore, predictable. Practical experience tends to justify the cautious "wait and see" approach that pervades the license exempt industry.

---

[2] An example is the regime for the 2.6 GHz band that allows conventional 3G systems (FDD) as well as WiMAX systems (TDD) to use this frequency band.

Finally, there is the cost and risk of designing and implementing generic spectrum sharing mechanisms. Absent regulatory requirements or industry standards, design criteria become a matter of choice rather than necessity and the risk of "non-performance" looms large. Taken together, these factors converge into a single question that technology vendors must answer: "Is there demand that justifies investing in generic spectrum sharing features?" So far, the answer appears to be negative. On the other hand, the implementation of vertical spectrum sharing between dissimilar systems is progressing well. In all cases, the sharing capabilities are required by regulatory regimes: C-band Wireless LANs,[3] C-band Broadband Wireless,[4] the TV White Space systems[5] and Ultra-Wide Band[6] are examples. Notably, the DFS function required for C-band wireless LANs is widely implemented and it has proven to be very successful, although amenable to improvement.

12.1.4 The Research Perspective

Research into the theoretical basis of wireless systems is broad and has a long history. A quick check on the web for the combination of the terms research and such terms as spectrum *sharing, CSMA/CA, MAC protocol, capacity and ISM band* give many tens or even hundreds of thousands of hits. Fuelled by economic importance as well as technical challenges, research of things wireless clearly is a fertile ground for interesting research papers and Ph.D. theses. In much of this work, the focus is on one or two aspects of a problem in wireless rigorous theory, and mathematical treatment frequently seems more important than the relationship with practice. There are many reasons for this state of affairs, but one that is relevant to the purpose of this book may well be the fact that practice is complex and hard to reduce to a set of simple, easily defined relationships or interactions. Some research focuses on the physical aspects like transmission, propagation and digital signal processing. Other research focuses on protocol issues, yet other work analyzes capacity issues in wireless networks; the list is long and varied. However, some[7] research papers address the intersection of the different aspects of wireless systems behavior. As the preceding chapters show, the interaction between propagation conditions, receiver parameters, and protocols gives rise to complex patterns of device behavior that tend to wreak havoc on performance and user expectations.

[3] The Frequency range is 5,150–5,725 MHz; it is available in most countries. See also Sect. 11.1 above.

[4] The frequency range is 3,650–3,700 MHz; it is available in the US only.

[5] See Sect. 11.2 above.

[6] See Chap. 8, Sect. 8.3.

[7] One example is the Rice Networks Group, led by Professor Ed Knightly.

As noted, there is no broad theory of "spectrum sharing" that brings together all relevant factors. In fact, an overarching theory may not be possible because of the large number of factors involved. The implications are twofold: firstly, the field will have to be broken up in different areas of research and, secondly, new directions of problem solving may have to be explored. Just as neural networks and genetic algorithms have made it possible to find solutions for intractable, complex problems, adaptive systems may prove useful in addressing the particular problems of spectrum sharing. Such an approach would likely be specific for a given sharing realm, e.g. for sharing among similar systems e.g. those operating in application specific commons spectrum, at least initially.

Another important line of research is that of Radio Resource Management. If we assume that the theory of wireless networking will provide much better insight into the factors that affect performance, the problems remain of how to express performance target, how to specify the elements of an RF plan, and how to quickly compute such plans. In other words, theory of operation needs a companion in a theory of control.

In summary, wireless networking promises to remain a rich field of fundamental as well as applied research. This book hopefully motivates some bright minds to engage themselves.

12.2 Evolution of Spectrum Sharing

The above perspectives show an increasing demand for spectrum resources and increasing attention to the issues related to sharing spectrum resources. How spectrum sharing will evolve is impossible to say: history shows that predicting the future remains difficult if not impossible. As indicated by the research perspective sketched above, advances in understanding are to be expected – as well as necessary.

Whatever the details, future developments of spectrum sharing theory as well as regulatory practise will shape up along two lines: sharing among similar systems and sharing among dissimilar systems.

The former is the more tractable of the two, and a rich basis of fundamental theory and practical solutions has been developed, fuelled to a large extent by the rapid growth of cellular communications.

The latter is the most intractable of the two and its future is, therefore, hard to assess. Operation in license exempt frequency bands that are open to all and sundry raises its own set of problems, not the least of which is equitable sharing of spectrum resources among dissimilar systems. A practical approach does not burden implementations unduly and keeps the door open to innovation. Regulatory authorities would most likely welcome progress on this subject looks into the possibility. Progress requires a generic spectrum utilization measure that could be used in developing simple rules for sharing spectrum among dissimilar technologies.

12.2.1 Constraints

Sharing calls for constraints – regardless its subject. In the case of spectrum sharing, these constraints are of two kinds: physical constraints and regulatory constraints. The former are largely immutable: the Shannon-Hartley theorem and the SNR wall are two examples of constraints that no technology or regulation can remove or ameliorate. This puts limits on the accuracy and reliability of spectrum sensing and renders it unfit for many vertical sharing scenarios. However, for horizontal sharing scenarios this objection does not apply.

Regulatory constraints are less permanent: they reflect the values society has put on certain ways of using certain pieces of RF spectrum. These values change with time and with the evolution of technology. Much of the current debate about spectrum management and spectrum scarcity is informed by these changes or, more precisely, by the perception of these changes and their importance. Changes in society's values attached to RF spectrum are largely driven by economics, a subject that falls outside the scope of this book. However, the ways in which the regulatory constraints evolve should be informed by an understanding of the technical factors that underlie spectrum usage and spectrum sharing. These factors include cost, capacity, efficiency, and fairness. They are discussed in more detail below, in the context of horizontal spectrum sharing, both between similar systems and between dissimilar systems.

12.2.2 Cost

A very basic fact is that the cost per bit of wireless communication is 100–1,000 times higher than the cost per bit of wired communications. This large difference is due to physical reasons, rather than commercial reasons. Another consideration is that the cost per bit of wireless transmission goes up with the factors of distance and transmission rate: both drive up the power as well as the required complexity. The difference in cost of an LTE base station and a IEEE 802.11 MIMO Access Point is instructive.[8]

These basic facts point towards the long-term evolution of communication systems in general and spectrum usage in particular: the coverage provided by wired communication systems will increase. Eventually, every office desk and every private room will have its multiple gigabit communications link. In the long term, even streets will become wired in some way. Given this increasingly fine-mazed wired coverage, wireless connectivity will become focused on short-range and high transmission rates. That, by its self, will take the pressure of spectrum demands and eliminate any chance of spectrum scarcity becoming a real issue – as opposed to an economic issue.

[8]One LTE market overview projects $11.400 M base station revenue for 225 M users. At an average of 50 users per base station, this equates to $2,500 per base station unit. This is >25 × times the cost of a Wi-Fi Access Point.

Clearly, wireless remains the only choice when mobility considerations prevent leveraging the dense wired infrastructure. Therefore, spectrum suitable for macro-mobility will remain in high demand. That being the case, one should expect continued pressure on regulators to make more spectrum available for macro-mobility applications.

12.2.3 Capacity

Wireless network capacity is ultimately limited by the available SNIR, not by sharing techniques such as listen-before-talk etiquettes. These serve primarily to avoid destructive interference at short distances.

Network capacity is largely a matter of data rate and operating range. The basic formula is:

$$c = \frac{\mathrm{f(pathloss)} \times \mathrm{R}}{\mathrm{area}} \tag{12.1}$$

in which f(pathloss) denotes the impact of pathloss on the propagation of wanted and unwanted signals and R stands for the average transmission rate, taking into account that interference between network devices limits the available SNIR.[9] Clearly, if the area served at a given transmission rate decreases, the capacity of the network increases. For example, packing more wireless LAN Access Points into a given space increases overall network capacity.[10]

If the network is homogenous, transmitter power and SNIR play no role: if all stations increase their power, the interference each node causes increases as well and this negates the possible benefit of the higher SNIR because the SIR dominates the achievable transmission rate. Conversely, reducing transmitter power may result in reduced SNR and therefore lower throughput.

The above simple formula does not apply in case of a heterogeneous deployment, in which different networks or network elements are owned and operated by different entities, but operate side by side in the same frequency band. Here, cranking up the power benefits the own network, but damages others. This consideration points towards a benefit of a homogeneous power limit for a shared frequency band.

12.2.4 Efficiency

If the factor capacity seems rather simple to determine and manipulate, efficiency is a far less straightforward factor. Whereas Capacity can be expressed as Mb/s/m-2,

---

[9] See Chap. 7, Sect. 7.2.5.

[10] This is true only if the nodes of the network do not "time share" the same channel.

Efficiency is determined by energy and throughput. In fact, by adding energy to Eq. 12.2, we get an expression for efficiency:

$$e = \frac{\text{f (pathloss)} \times \text{R}}{\text{area} \times \text{energy}} \tag{12.2}$$

For a given pathloss, e is determined by the data rate divided by the area and the energy of the transmission. Written a different way:

$$e = \frac{\text{C} \times \text{R}}{\text{p} \times m^2 \times s^2} \tag{12.3}$$

in which C is a constant for the applicable pathloss model, R is the data rate and p is the RF power used. This clarifies that time and coverage are the most important factors that determine the efficiency of wireless communications.

Maximizing efficiency requires minimizing time. However, minimizing time requires more power. One might think that, because the power factor has linear impact, increasing it is worth the gain in time reduction. However, p also determines the SIR seen by receivers; and since data rate increases as the logarithm of the SIR, the decrease in transmission time follow the log of the power increase. This observation confirms the results given in Chap. 7, Sect. 7.2.5 on spectrum re-use: there is an optimum data transmission rate that maximizes throughput as well as efficiency. In the case of wireless LANs, this maximum is reached with the use of QPSK modulation.[11]

12.2.5 Fairness

Fairness is probably the most difficult factor to analyse, if only because of its social background. Ignoring that, one has to find an expression or unit of measure for fairness in wireless systems. The subject has a considerable history in computing and in the TCP world. A key paper dates from 1984.[12] It divides the question of fairness in two parts: the selection of an appropriate metric and the definition of a relationship that expresses a quantitative assessment of fairness, i.e. a "fairness index" (using a chosen metric). The paper is only concerned with finding that relationship. It mentions four desirable properties of the fairness index: size independence, scale and metric independence, boundedness and continuity. It defines fairness as:

$$I_f = \frac{\left(\sum x_i\right)^2}{\left(n \sum x_i\right)^2} \tag{12.4}$$

in which x_i is the share that i gets. How that share is measured is not addressed.

[11] See Chap. 7, Sect. 7.2.5.

[12] See Jain et al. [70].

Another way of looking at fair sharing of spectrum resources is related to the efficiency of spectrum use. This is known as proportional fairness and takes into account the relative demand on resources of users in allocating those resources. Inefficient users get less access to the available resources, but they are never starved. A similar approach was developed by one of the authors in a study for the Wi-Fi Alliance. The basic concept of this approach is that overall value of the spectrum is maximized when the efficiency of its usage is maximized. Thus, transmitters that successfully deliver to receivers get easier access to the channel than transmitters that are less successful, e.g. because of retransmissions. In a network with multiple RF channels, this policy will force unsuccessful transmitters to look for a better channel, rather than consuming resources that would be better used by others. If applied in a homogeneous network, this policy could be an integral part of a distributed Radio Resource Management system.

12.3 A Generic Sharing Metric: The Transmission Unit

Sharing calls not only for constraints, it also requires a metric that can be used to express amounts of what is shared and in what way. Taking into account the preceding discussion of capacity, efficiency and fairness, the following proposes such a metric and some rules for using it in different contexts.

12.3.1 Background

The measure of transmission resources to be allocated to transmitters can be expressed in many ways, e.g. time units (=channel occupation), throughput (e.g. frames per second), etc. All have the disadvantage of not being adaptable to local, possibly temporary conditions. A generic metric is needed that does not suffer from this disadvantage.

Such a metric might take the form of a "currency," the units of which can be defined, allocated and exchanged as required for a given context or frequency domain. The unit of such a currency could be defined as a combination of resources allocated to a transmitter at a given point in time; once the unit is exhausted, the transmitter has to relinquish the channel so that another transmitter can take over. Such a unit could be called a "transmission unit."

The transmission unit is well suited for the "good behavior is rewarded" model of proportional fairness: efficiency results in more ineffectiveness results in less transmission units. In any case, these transmission units are limited in one or more ways. In case the spectrum concerned is dedicated to some application or system, the measure of the transmission unit can be simple – to some extent, it is arbitrary and its definition can be left to the system's designer or spectrum owner. However, a general purpose definition has to be independent of application- or technology specifics.

12.3.2 Definition

In case of a spectrum commons which admits a variety of systems and technologies, expression of the transmission unit in physical terms is necessary. To have a technology neutral definition of a transmission unit, one needs to address all relevant dimensions: frequency, power, time, and space. A fully open commons would be well served with the definition of transmission unit as a dimensionless number

$$U_{tx} = \frac{f \times e}{\chi} \tag{12.5}$$

in which f stands for frequency, e stands for energy (=power × time) and χ stands for (interference) space or footprint.

These elements are all expressed as ratios relative to reference values that are unique for the frequency band concerned. An example based on the 5 GHz regulatory limits for indoor wireless LANs may make this clear. The amount of spectrum concerned is 455 MHz, the power limit is assumed to be 200 mW. Maximum duration of a transmission should be in the range of the channel coherence time, e.g. 2 ms, and the space factor reference can be set at 360°. If we assume the default channel width is 20 MHz, the reference value- U_{max} -of a transmission unit would be $1 \times 1 \times 1 \times 1 = 1$. The actual consumption of spectrum resources could be measured against this, as the examples below illustrate.

(a) A device using 20 MHz and 50 mW, for 1 ms per transmission using omni directional antenna would have a $U_{tx} = 1 \times 50/200 \times ½$ x = 0.125 which is eight times less than U_{max}. This margin could be leveraged to extend transmission time.
(b) A device using a 40 MHz channel width, 200 mW power output and an omni-antenna would consume $2 \times 1 \times 1 \times 1 = 2$. Since this would exceed the maximum value 1, one or more of the parameters has to be reduced.

Applying the same considerations to a 20 MHz, 1 W outdoor device with a 30° antenna in the 255 MHz available for outdoor use at 1W EIRP and assuming a 2 ms basic time unit would give $X = 1 \times 1 \times T \times 1/12 = .0833$ T. This would allow such a device to transmit up to $4 \times 12 = 48$ ms without spectrum sensing.

Another example is a 60 GHz device with a channel width of 100 MHz, a 1° beamwidth operating in 2 GHz of license exempt spectrum. Assuming a default channel width of 100 MHz and the same reference values for power, time and antenna angle gives 1/360 as actual resource consumption. This margin could be used to increase transmission time up to 2×360 ms = 720 ms. Thus, it would be allowed to transmit for 720 ms each time it detects a free channel. A small reduction of the power output would allow it to transmit for 1 s per second, interspersed with short sensing periods.

12.3.3 Application

Although the Transmission Unit can readily be used to express resource utilization, it says nothing about rules for allocating resources. Rules for allocating Transmission Units can be simple and non discriminatory, for example:

use one unit, then check the channel[13] *and if free use another unit.*

Such a rule seems easy to implement and fair, but it is not: inefficient transmitters get as much resources to spend as more efficient ones; this wastes spectrum resources and, therefore, some form of proportional fairness constraint must be added, for example:

use one unit, *if not successful, wait*, then check the channel and if free, use another unit.

The "if not successful" restrains unproductive use of spectrum resources and encourages designers to find ways to make their systems efficient.

12.4 Conclusion

The preceding example of the Transmission Unit makes clear that there is ample opportunity for inventing new ways to deal with the sharing of spectrum for various purposes. However, the benefits of medium access rules like the above cannot be realized unless accommodated by the applicable regulations for a given frequency band. Regulatory authorities[14] may have to adjust their methods of specifying regulatory limits and technical requirements in order to accommodate Transmission Unit based spectrum sharing in their rulemaking.

12.5 Wrap-Up

The use of wireless technologies is spreading to all aspects of our daily life and to every corner of society. The rate and scale of the adaptation of these technologies is phenomenal and there is no reason to expect that this will change. In all of these technologies, the ubiquitous availability of free spectrum plays a key role. However, that spectrum has to be shared – simply because there is not enough spectrum to give every system and every application its own spectrum. With time, the challenges of capacity limitations and pervasive interference will become increasingly real. Wireless

[13] Checking the channel is a complex subject by itself; rules for channel checking may be context specific.

[14] The FCC has recognized the value of directional antennas and has adjusted its Part 15 rules accordingly. In some cases, it allows unlimited antenna gain.

management, especially Radio Resource Management, is therefore one of the most important tools in efficiently using shared spectrum resources.

This book is intended to help addressing these challenges by providing theoretical insight, as well as analysis of regulatory policy and current practice. All are relevant to the continued success of commodity wireless technology. Although the subject of spectrum sharing is complex and its variables many, a few broad conclusions can be drawn.

First of all, spectrum scarcity is a convenient catch phrase rather than a reality. The capacity of wireless systems is essentially a matter of physical scale: at small scales, that capacity is very large and in practice nearly unlimited.

Secondly, because of the diversity of wireless technologies there is no best way of sharing RF spectrum: the interactions between different types of radio technologies as well as environmental variations introduce asymmetries that cannot be eliminated.

Thirdly, physical and technical constraints limit the accuracy of spectrum sensing and geo-location. Deterministic control of vertical spectrum sharing is, therefore, likely to prove feasible only in special cases,[15] rather than in general.

Therefore, the search for simple spectrum sharing rules and technical solutions that are neutral with regard to technologies and application is misguided. An example is the listen-before-talk concept, which seems destined to become a pet solution of many even though its efficacy varies widely and unpredictably. Constraints on spectrum resource utilization, phrased in terms of frequency, energy and space coupled with proportional fairness rules, promise to leave many avenues open for innovation and prevent "winner-take-all" technologies from becoming dominant.

As is the case with most if not all human endeavours, there is no free lunch in spectrum sharing matters: efficiency and performance increase, but only with loss of generality.[16] The observation that specificity enables performance points towards the need for reconsidering different categories of spectrum and spectrum use. ISM bands have proven great breeding grounds for new technologies, but broad, sustained use of these technologies is best served by providing spectrum for a given class or type of technology – a technology specific or application specific commons approach.

Regardless of how regulatory authorities decide to proceed with spectrum allocation, assignment and management, spectrum sharing theory and practice will remain a rich field of research, as well as practical innovation for the benefit of all.

[15] i.e. DFS for wireless LANs and DAA for Ultra-Wide-Band.

[16] The IEEE 802.11 standard is a good example: its most recent extensions provide a throughput of hundreds of Mb/s, but its spectrum sharing behaviour is far less general and far less benign than that of the early base standard.

Appendix A: Ofcom's Technical Studies

Ofcom's Study of Politeness Protocols

In 2005, Ofcom started with a study that aimed at understanding the possible benefits of politeness protocols in the context of license exempt spectrum sharing.[1] One of the assumptions was that improved sharing protocols might provide improved spectral efficiency and/or allow for higher power levels without causing "excessive interference." The scope of the study, as formulated by the research team, was to investigate polite protocols that are open, simple, economically feasible and that allow heterogeneous systems with a wide range of data rates and traffic models to operate in common spectrum with minimal interference.

The study looked into the properties and effectiveness of a number of methods and protocols, not all of them existing at the time the study was performed.

The study also considered two sharing regimes: homogeneous and heterogeneous – both are relevant because large differences could point to the need for different regulatory measures. The study also considered two deployment scenarios: a 2.4 GHz scenario using contention-based protocols and a 5 GHz scenario using time sharing (synchronized and unsynchronized) protocols. The study considered the methods and protocols listed in Table A.1.

[1] See http://www.ofcom.org.uk/research/technology/research/exempt/polprot/.

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1, © Springer Science+Business Media B.V. 2011

Table A.1 Protocols used in the Ofcom politeness protocols study

Method/protocol	Cooperative?	Example
Intelligent Duty Cycle	Yes	Various
Unsynchronized Time Sharing	Yes	Novel
Synchronized Time Sharing	Yes	Novel
Centralized Polling	No	802.11
Distributed Polling	No	Unknown
Time Spread Multiple-Access	Yes	Unknown
Receiver Beacons	Yes	Novel
CSMA/CA	Yes	802.11, HiperLAN1, Zigbee
RTS/CTS	Yes	802.11
Not Clear To Send (NCTS)[2]	Yes	Novel
Transmit as Receiver[3] Beacons (TARB)	Yes	Novel
Free Channel Search	Yes	DECT
Dynamic Channel Assignment	No	Unknown

The summary of the results[4] given here is not complete, but focuses on the main conclusions:

- CSMA/CA is a good general purpose protocol, but it could be improved by receiver feedback – e.g. the TARB model – notably in the case of indoor scenarios.
- Time sharing methods work well with heterogeneous systems, as would an intelligent duty cycle system, but it is recognized that these approaches may cause problem w.r.t. response time and/or QoS.

Whether these conclusions, given the considerable "novelty" content in the protocols considered, will prove an adequate basis for policy development remains to be seen.

Ofcom's Consultation on Spectrum Commons Classes

In its consultation on "Spectrum Commons for License Exemption," Ofcom noted that it believes that Application Specific Allocations for license exempt devices results in fragmented and inefficient use of spectrum. However, Ofcom also noted that the benefits of license exempt operations are maximized when technologies with similar operating parameters are grouped together in a given Spectrum Commons Class.[5] The definition of such a Class is given in terms of an interference potential, captured by a single variable, the Interference Indicator. In this perspective, the Interference Indicator only has relative value, not absolute value.

[2] NCTS allows receivers to send a "blocking signal" that avoids interference from unwanted transmitters – this is not unlike the Clear to Send element of the IEEE 802.11 RTS/CTS protocol.

[3] TARB will prevent transmission if the intended receiver is not detected.

[4] See http://stakeholders.ofcom.org.uk/binaries/research/technology-research/summary2.pdf.

[5] See http://stakeholders.ofcom.org.uk/binaries/consultations/scc/summary/SpecCommonsClasses.pdf

The Interference Indicator is constructed from four parameters types: bandwidth, time, coverage, and density. The bandwidth parameter gives the fraction of the available bandwidth used, the time parameter gives the duty cycle of the interferer, and the coverage parameter takes into account both the range of the interfering signal, the antenna beamwidth and the density of the potential interferers. The range of the interfering signal is determined by the EIRP of the interferer and the receiver's sensitivity for interference. For the latter, Ofcom proposed a threshold of –80 dBm/MHz – this is 34 dB above the thermal receiver noise floor. Coverage is expressed in km^2.

Interferer density was proposed to be based on a "busy, yet realistic scenario" for a given technology. Density is expressed in device/km^2. The following Table A.2 gives some results given in the Consultation for existing technologies.

Table A.2 Interference Indicator values for selected technologies

Technology	Bandwidth factor	Duty cycle	Coverage (km^2)	Density (units/km^2)	Interference indicator
RFID	.002	.1	.5	234.8	.028
IEEE 802.11b	.263	.012	.003	15,000	.164
Bluetooth HID	.012	.25	.003	12,500	.105
Home Automation	.0012	.0001	.43	20,000	.0014

These results clearly depend on the assumptions for coverage and device density and, therefore, it is not clear that these results are useful in any other but an indicative sense. Some objections to the parameters proposed in this Consultation are:

- The bandwidth factor ignores the effect of asymmetry of bandwidth: narrow band devices affect wideband victims at much larger distances than vice versa.
- The duty cycle factor does not take into account frequency of access, duration of access, and probability of access.
- The interference threshold that determines the coverage area favors narrow band devices.
- The density factor is highly subjective and will vary much with location, time of day, and type of use.

The above results – which are almost equal for highly incompatible systems like Wi-Fi and Bluetooth – expose the weakness of the Interference Indicator approach. Although different from the FCC's Interference Temperature concept, the Interference Indicator concept is also based on the idea that the impact of dissimilar systems and technologies can be compared given adequate abstraction of the method of measurement.

Appendix B: DFS – Background and Compliance

As described in Chap. 9, "Dynamic Frequency Selection" or DFS became the enabling factor in the allocation of 5 GHz spectrum to wireless LANs and similar devices. This Appendix adds more details and summarizes the regulatory requirements and the compliance criteria that were developed to assess actual compliance of equipment.

Background: Radar Detection and Avoidance

The concept of DFS is simple: if a wireless LAN detects a radar signal, it has to move its operating channel to another frequency so that its transmissions will not interfere with the radar's operation. That simple concept was the basis for the DFS requirements and the test patterns developed in ITU-R Recommendation M.1652 and its descendants. The practice of DFS is far from simple: radar signals vary in the spatial, the frequency and the time domain, and detection is affected by the wireless LANs own operations. Conversely, the radar's behavior, the absolute distance and the pathloss between wireless LAN and radar affect the interference. The following looks into the factors that affect radar detection and wireless LAN, causing interference.

Radar Signal Properties

Radar Types and Radiation Patterns

In order to understand how DFS has to work and how it can be tested, one has to know the key radar parameters: power output, antenna gain, rotation/scan rates, pulse modes and rates, and the I/N protection margin.

Many radars which operate in the 5 GHz band are used for weather analysis, navigation and/or military target acquisition purposes. There are fixed and mobile ground-based radars of this type, as well as shipborne radars. These are high power systems with power outputs in the range of 50–150 kW. Antenna gain is in the range of 40–50 dB and rotation speed is typically low, e.g. 6 rpm. Pulse duration is in the range of .5–5 μ and repetition rates are commensurate with the long-range – 250 pps for older radars. Newer designs use faster but multiple pulse rates, which avoids the range ambiguity associated with a single pulse rate. 5 GHz meteorological radars are deployed world-wide. They are used for a variety of purposes, including not only storm warning, but also rain measurement, wind direction, etc. Doppler techniques and advanced signal processing allow precise, volumetric data to be collected over large areas. The output power and pulse patterns are similar to long-range radars, but the beam shape (narrow pencil beam) and scanning strategies (helical) are different – the same spot in space may be scanned only once in 5–10 min. Current meteorological radars make use of pulse widths between 0.5 and 2.5 μs. The typical PRF used ranges from 250 to 1,200 pps and are frequently staggered. Some type of "met" radars use noise calibration of the radar receiver, with the transmitter turned off at antenna elevation angle typically between 15° and 60°. At these elevations, these radars do not see ground-based wireless LANs and vice versa and, therefore, these noise calibration periods do not increase the probability of wireless LAN interference – contrary to claims of weather radar community.

In addition to these large fixed radars, there are some mobile tactical radars that operate in the 5 GHz band. They are typically used for local air defense and battlefield surveillance. Due to size restrictions, antenna gain is usually lower, e.g. 30–36 dB, they use lower power transmitters, higher pulse rates, and high antenna rotation speeds. Advanced digital signal processing is employed to counter interference and jamming.

Radar detection by wireless LAN devices poses a complex technical problem involving – in the case of military radars – some secret parameters. However, civilian technology tends to be ahead of military technology and not just in case of cellular communication systems. Using material on civilian and military radars available on the internet, a coherent picture of the relevant aspects of radar emissions can be put together.

The Detection Threshold

The detection threshold for DFS was determined by simulations[6] performed by experts of the NTIA during the preparations for WRC 2003. These simulations included a large number of parameters that described the density distribution and location of wireless LANs in a hypothetical city on the coast. This allowed maritime

[6] See Report 09–461 available from http://www.ntia.doc.gov/osmhome/reports.html.

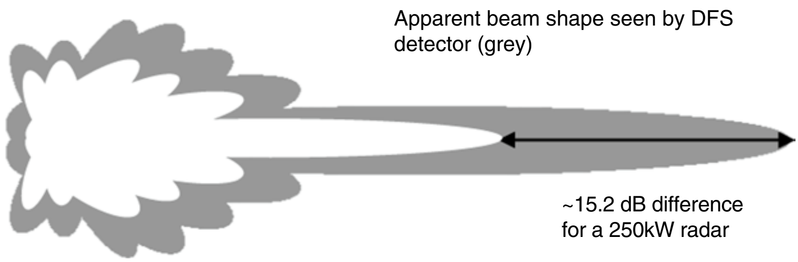

Fig. B.1 Apparent beam shape differences

radars to be accounted for as well. The resulting figures of –62 dBm for indoor wireless LANs and –64 dBm for outdoor wireless LANs became the widely agreed basis for the radar-wireless LAN co-existence regime.

Antenna Gain and Apparent Radar Beam Shape

Antenna gain and directionality are key elements of radar systems, regardless of the use of physical or electronic control over the direction and beam shape or pattern of emission. In practice, antenna size is the main determinant of antenna main beam gain and sidelobe gain. More reflector surface means more gain and, with a reduced illumination factor, it allows for lower sidelobes. In practice, antenna size is compromise involving many factors including size, cost, wind loading, etc.

From a DFS detector point of view, a radar beam is usually broad and has a few sidelobes, but that was not taken into account in setting the DFS compliance criteria. These are based on a nominal beamwidth of about 1°. Nominal radar beam shape is given in terms of the half power beamwidth for the main beam and major sidelobes. This nominal figure is useful to determine the spatial resolution of a radar at extreme range or minimal target size, but it says very little about the signature of a radar signal as seen by a DFS detector that is located rather close to a radar.

In general, the link budget for the detection of a radar by a DFS detector is much greater than the link budget that determines wireless LAN interference seen by the radar. This is shown in the figure below, for a high power radar system of 250 kW output power and a wireless LAN of 200 mW radiated output.

This link budget difference is determined by the EIRP of the transmitters and the thresholds of the receivers and not by parameters of antenna gain, pathloss, etc.; these affect the actual distances involved. The radar horizon determines if a radar is visible to a DFS detector and vice versa. Due to the curvature of the earth and terrain undulations, tall buildings, etc., the actual radar horizon is typically much less than 50 miles away and sometimes much less. Within this limited horizon, the radar

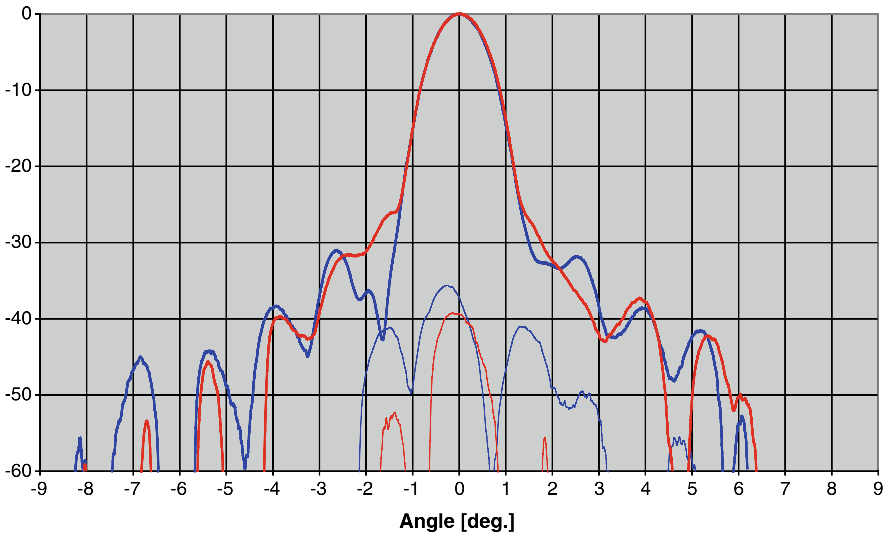

Fig. B.2 Beamshape of a modern weather radar.[7]

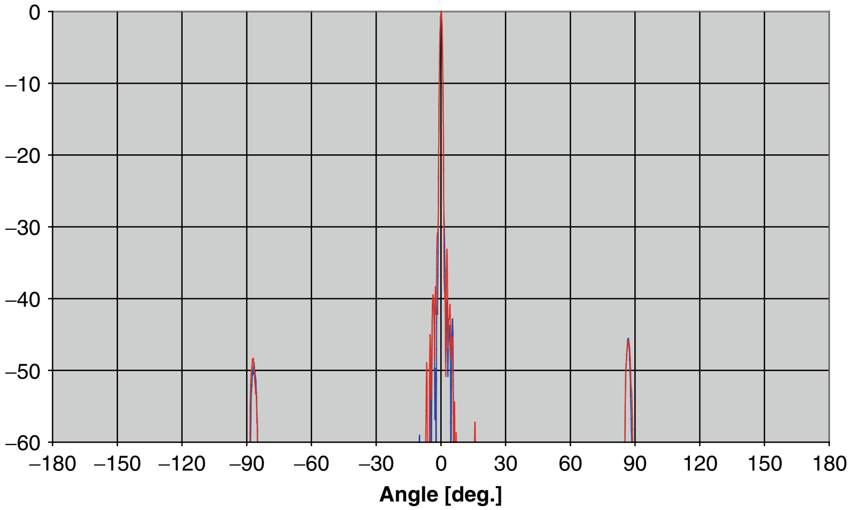

Fig. B.3 Main sidelobes of a modern weather radar in dBr[8]

[7] Reproduced with kind permission of the Vaisala company.

[8] Idem.

signal strength is fairly high and, therefore, well above the DFS threshold, even for an indoor wireless LAN.

Assuming a typical terrain with pathloss varying between 6 dB/octave and 12 dB/octave over a distance of 50 km, the residual signal strength of a 250 kW radar is (84 + 46)–47–105.7 = −22.3 dBm. This exceeds the detection threshold by some 40 dB and, therefore, the effective beamwidth of the radar seen by a DFS detector at this distance is the –40 dB beamwidth. This is much wider than the nominal beamwidth, even for an advanced pencil beam antenna designed specifically for high resolution weather radar use.

As the Fig. B.2 below shows, the –40 dB beam width is nearly 6° wide – that is about seven times as much as the nominal beamwidth.

Radar Antenna Sidelobes

All radar antennas have sidelobes, even the high quality designs of the Vaisala company of Finland. Figure B.3 shows that, in case of the latter, sidelobes of up to 50 dB below the main beam are visible at right angles to the main beam.

However, many radars produce larger sidelobes. Older designs are less benign and, notably in the case of smaller antennas, may show sidelobes up to 20 dB below the main beam. In practice, a safe assumption is that major sidelobes will range from 30 to 50 dB below maximum gain.

The impact of such sidelobes is that, depending on the distance to the radar, DFS detectors will see multiple “images” of a radar at different angles, with the number of such angles depending on the power of the radar and its antenna pattern. Given that the link budget difference for detection and interference is independent of antenna gain, these sidelobes do not correspond to potential interference threats and they do not contribute to radar detection.

Observed Burst Length

The rotation and scanning motions of radar antennas limit the time a given point in space is illuminated by the radar beam. To yield useful information for the radar receiver, this illumination has to deliver a certain amount of energy. This amount of energy is determined not only by the transmitter’s power output and antenna gain, but by the scan rate of the antenna, the pulse repetition rate of the transmitter and the pulse width. For a given resolution, the product of the dwell time, power and pulse widths has to reach a certain value. Reducing one factor may require increasing another. Given the need to deliver sufficient energy to any point of interest, detection of radar signals by DFS detectors appears assured.

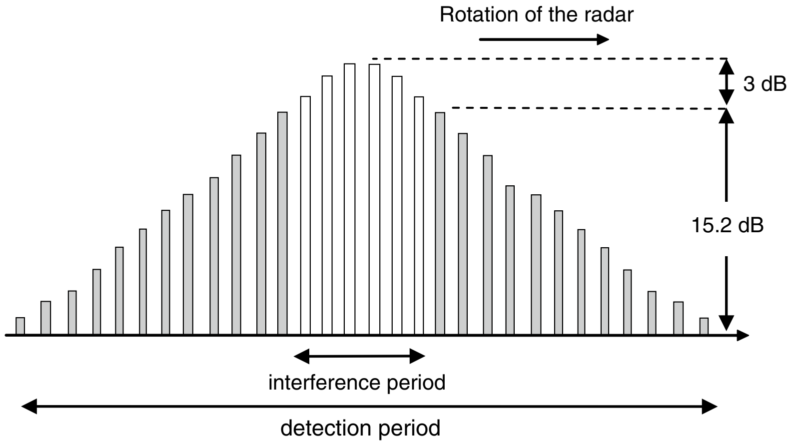

Fig. B.4 Timing difference between radar detection and wireless LAN interference

The observed burst length is given by the pulse repetition rate (PPR) divided by the dwell time of the beam: PPS*°BW/(6*RPM). A typical weather radar has a burst length of 900*.9/36 = 22.5 pulses – at the nominal beamwidth BW. As explained in more detail below, a key factor in radar detection is the number of radar pulses seen by a detector: more pulses mean that the detector can more clearly distinguish between actual radar pulses and spurious effects – even in the presence of wireless LAN traffic.

The benefit of the large margin between the radar signal and the DFS threshold is that the number of pulses seen by a detector is large. Although Fig. B.4: Timing difference between radar detection and wireless LAN interference is not exact, it shows the effect: as the radar beam moves over the detector, detection occurs well ahead of the point where the wireless LAN signal could cause interference (if DFS were not active).

This figure is representative for a detector at 50 km distance. At shorter distances, the difference increases to the point where antenna sidelobes also become detectable. At larger distances, the difference decreases; but due to the horizon cut-off, this has negligible impact on radar detectability.

An analysis covering nearly all weather radars in Europe[9] showed that, except for a few old radars only used for storm warning, the number of pulses seen by the DFS detector would be in the range of 43–145 pulses per "radar sweep." This is adequate to assure a very high detection probability indeed.

[9] See ECC Report 140, source ERO.

Pulse Detection Statistics

In practice, the DFS detection efficiency is determined by the statistics of the pulse patterns (intervals and burst length) and by operational conditions such as the wireless LAN busy level. Implementation plays a limited role. In general, the detection efficiency of a detector follows a cumulative binomial distribution function: it determines how the detection probability p for individual pulses affects the overall burst detection probability P for n such pulses, given a minimum of k pulses to be detected.

$$P_{(k\ out\ of\ n)} = \frac{n!}{k!(n-k)}(p^k)(q^{(n-k)}) \tag{B.1}$$

where k = the false alarm threshold, n = the number of pulses in a burst, p = the detection probability for a single pulse (or the listen/talk ratio), q = the probability that a pulse will not be detected.

In this case, q = 1–p because the detection and non-detection are each other's complement.

This formula ignores the aspect of how to relate each detected pulse to the other pulses – that is an implementation matter. The table below shows how detection probability varies with the number of pulses per burst and with the detection probability for individual pulses – in this case, for a comfortable false alarm threshold of 4 pulses.

Figure B.5 shows that there is a comfortable margin for reliable radar detection for all radars – with the possible exception of fast rotating radars with slow pulse rates. These radars are few and far between – if only because the spatial resolution achievable is meager.

Staggered pulse patterns affect the detection probability in different ways: frame staggering results in bursts of frames at some constant spacing and, therefore, detection probability is the same as for constant pulse rates. Pulse staggering requires more pulses because to the reduced redundancy in the pulse train: a 6 pulse false alarm threshold (=50% more) provides additional redundancy to compensate for this. Given the high pulse rates of staggered PRF radars, this 50% is easily reached.

In summary, whereas an analysis based on nominal radar data suggests difficulties in detecting certain radars because of the short burst lengths seen by a DFS detector, the actual burst lengths observed from real radars show ample margin and therefore high detection probabilities. In fact, radar operators have considerable influence on the detectability of their systems: higher PRFs mean more pulses and therefore increased detectability.

Note that the same detection margin applies to low-power radars, such as small ship borne radars or mobile battlefield radars. The Saab Giraffe radar is a case in point: because of its small antenna, its nominal beamwidth is wider than that of e.g. a weather radar. In combination with the high pulse rate of such radars, the number

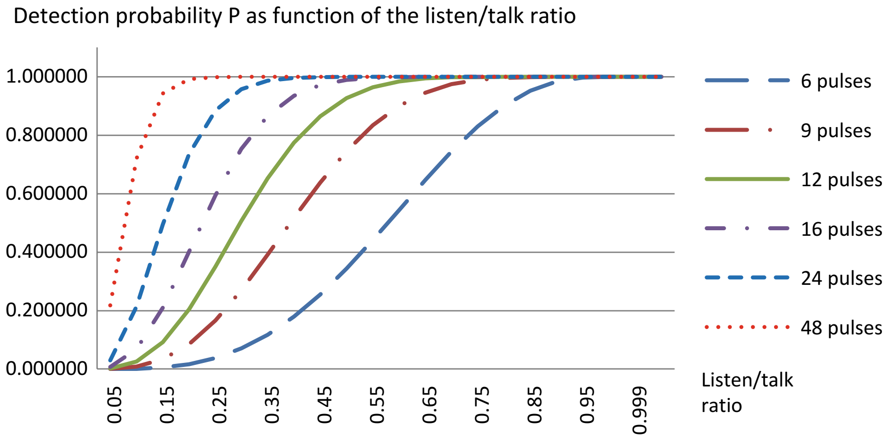

Fig. B.5 Radar Detection probability under different conditions

of pulses seen by such radars is more than adequate for reliable DFS detection. However, these considerations have not informed the DFS compliance criteria – there are all based on nominal radar parameters.

Pulse Shape and Spectrum

Radar pulses tend to be short: from .5 to 5 μs, and rising and falling edges tend to be commensurate: 1–10%. Steep slopes give rise to a wide frequency spectrum of short duration.

To assess the DFS detection performance at the frequency of interest, the radar's spectral power must be evaluated against the wireless LAN's DFS threshold. In general, the RF spectral power depends on the pulse rise time and fall time.[10] However, for the purposes of this document, one must look at practical data points.

Table B.1 shows, for a radar with a typical duty cycle of .1%, how the spurious emissions vary with pulse rise time (the fall time is assumed to be roughly the same).

Clearly, the best case from a wireless LAN point of view is the 10% rise/fall time: the RF power envelop falls off quickly to low levels. The 2% rise/fall pattern is closer to that of the ideal square pulse that needs to be considered. The last column gives a set of *conservative* working assumptions – from the viewpoint of a DFS designer – for a typical modern radar system.

Short rise times, necessary for certain radar modes of operation, cause a wider spectrum and, therefore, trigger radar detection over a wider spectrum than the

[10] See e.g. ITU-R Recommendation SM.1541 and ITU-R Report F.1097.

Table B.1 Radar RF power roll-off in the spurious domain

Δf (MHz)	ΔP, ideal square pulse (dB)	ΔP, 2% rise/fall time (dB)	ΔP, 10% rise/fall time (dB)	Working assumption (dB)
20	–33	–38	–45	–42
30	–39	–45	–55	–50
40	–42	–50	–65	–55
80	–50	–65	–80	–70

PRF = 1 kHz, pulse width is 1 μs

wireless LAN bandwidth. The implication is that wireless LANs will detect radars at frequencies at which they may not cause interference. Increased wireless LAN bandwidth – like used in MIMO systems – aggravates this difference – to the disadvantage of the wireless LANs.

As radar pulses travel along earth's surface, they encounter reflections and absorptions of many different kinds. Although one would expect this to affect the pulse shape, practice shows this to be not that case. The simplest explanation is that the hard scattering of the beam of edges of objects like buildings is very lossy and, therefore, the residual energy that travels with the main pulse is low enough so that the pulse shape is mostly preserved – at least in the time domain. This is important for detector design.

Radar Unwanted Emissions

Whereas the above pulse spectrum is considered part of the necessary bandwidth of a radar – necessary in terms of the radar's mission – the unwanted emissions that extend further out from the fundamental are not necessary, but they are a consequence of design and implementation choices.

Radar spectra vary with the design of the transmitter amplifier – klystrons tend to be cleaner than magnetrons and filters can be applied to reduce the "skirts" without affecting the necessary bandwidth much at all. The importance of understanding these spectra lies in the impact on the false positive detection rate of a DFS detector. The "dirty" spectrum of a magnetron radar may show peaks of –30 dBc at 10s or even 100s of MHz away from the fundamental. If we assume that the effective bandwidth of a DFS detector is 3 MHz, the relevant spurious for such a radar appear to extend contiguously up to 40 MHz away from the fundamental and re-appear at up to 200 MHz away. For a klystron-based radar, the spurious are typically more benign and up to 15 dB lower.

The apparent signal level seen by the detector varies with its bandwidth.[11] At an effective bandwidth of 20 MHz or more, the apparent strength of the spurious

[11] See http://www.its.bldrdoc.gov/pub/ntia-rpt/05-420/05-420.pdf.

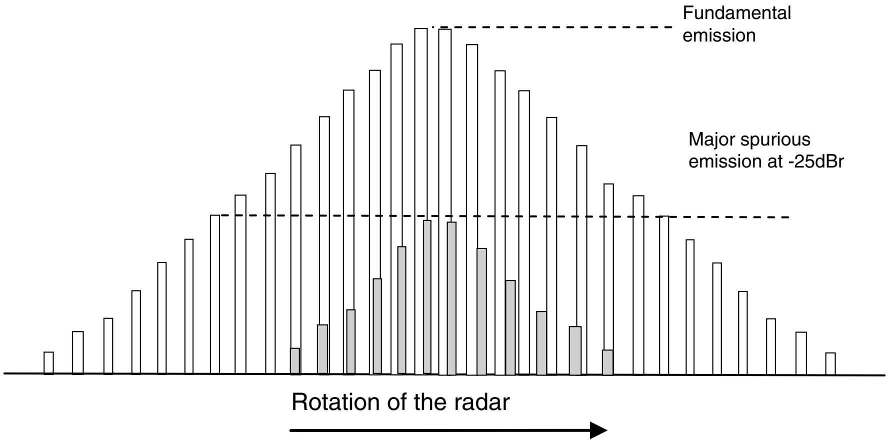

Fig. B.6 False DFS detection on a radar spurious peak of –25 dBr

emissions is well above the DFS threshold, even at considerable distances. For example, a major spurious emission peak at –25 dBr will be detected at up to 21.7 km distance from a 250 kW radar. See Figure B.6. It would be detected as a 12 pulse burst – more than enough to trigger the DFS channel blocking process – quite unnecessarily.

The unwanted radar emissions may not need to trigger DFS detection, but are not easily separated from the "real" radar emission. The first factor to look into is the effective DFS detector bandwidth – keeping that to 1 MHz should help to cut the false positive detection rate. A second factor to look into, notably for outdoor systems, is to correlate radar signatures over the whole 5 GHz band – not just the sub-bands in which transmission is allowed. If two radar signatures differ only in strength – and therefore in pulse count – the weaker signal is likely to be a spurious component rather than the main emission.

However, none of the detections caused by outlying spurious peaks would correspond to real interference threats because the radar is not sensitive at those frequencies. The implication is that a DFS detector could see a number of "occupied channels," whereas there is only one occupied channel in reality. This is not true for the emission levels close to the fundamental – the threat of these necessary bandwidth components must be verified against the actual frequency distance over which these spurious detections occur.

The spectrum of a typical C-band radar at –40 dBr is 40–60 MHz wide. A DFS detector would see this level as a valid radar signature at up to 20+ km away – depending on the power of the radar. Therefore, the radar's spectrum will prevent wireless LAN operations in an area of some 40–60 km in diameter over a width of 60 MHz or more – depending on the quality of the radar transmitter amplifier. This provides an additional degree of protection for radar systems.

Radar Scan Patterns

Radar scan patterns vary with the application and with advancing technical development. Little is known about military radars. Many are still based on the basic model of a rotating antenna, combined with a vertical fan beam or vertical scanning. Some combine search and tracking functions. Some more recent designs combine electronic beam steering with frequency hopping, which improves jamming resistance and reduces detectability. The most advanced civilian radars are weather radars. These combine staggered PRFs, multiple polarizations, and helical scan patterns. The latter has consequences for DFS operations. The example in Fig. B.7 shows that such a scan pattern may visit the horizon only for a few revolutions per scan cycle. DFS detection procedures must take this into account.

It should be noted that the detectability of these radars depends not only on the distance, but even more so on the elevation of the radar antenna. The strength of the radar's signal drops off as a function of the antenna directivity. See the example above of a state of the art weather radar antenna. At 4° elevation, the antenna gain has dropped by some 40 dB and, therefore, the detection distance drops to some 21 km. At 6° elevation, detection is virtually impossible – and so is interference from the wireless LAN into the radar at this elevation and higher.

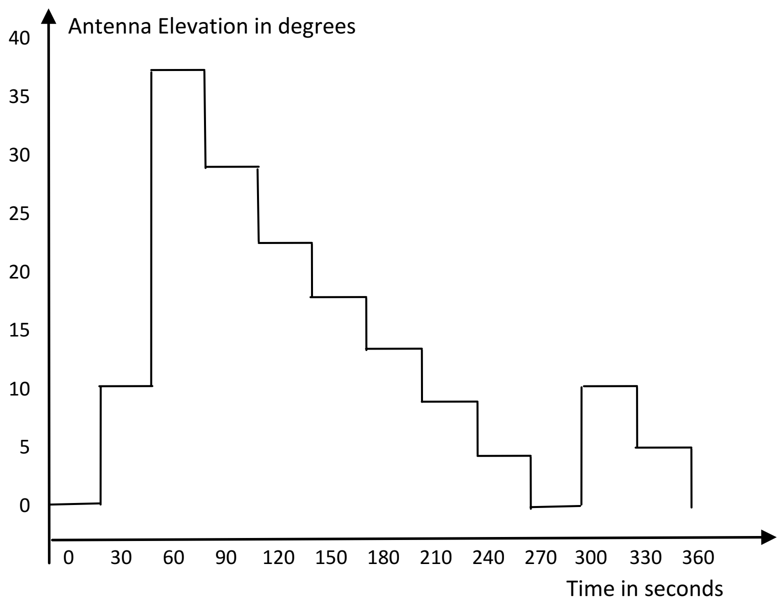

Fig. B.7 Typical weather radar scan cycle – elevation versus time

DFS: Regulatory Requirements

ITU-R Recommendation M.1652

In 2003, after a preparation of many years, the International Telecommunications Union (ITU), at its World Radio Conference 2003 (WRC-2003), agreed on a new frequency allocation on a co-primary basis to the mobile service for the implementation of wireless access systems, including radio local area networks (WAS/RLANs) systems, noting however that in the bands 5,250–5,350 MHz and 5,470–5725 MHz, stations in the mobile service should not claim protection from radiodetermination services. This was subject to technical and regulatory provisions included in the radio regulations, given in Resolution 229 [] (WRC-03) that makes the Annex 1 of ITU-R Recommendation M.1652 mandatory. This includes specific provisions to protect the incumbent systems; including military and weather radars.

The Recommendation spells out how wireless LANs have to behave in the frequency bands used by radars and it provides criteria for detection and RLAN response behavior. It has the force of international law: in case of interference conflicts between countries, proof of compliance with the "DFS Recommendation" is required. For products, national – or regional – rules and compliance criteria determine market access and/or the right to use equipment that is capable of operating in these bands. All of these are based on M.1652 which lists many considerations that led to the restrictions imposed on RLANs so as to assure a stable sharing regime. The main elements of the requirements of the Recommendation are given in quotation below.

2.1 Detection requirements

The DFS mechanism should be able to detect interference signals above a minimum DFS detection threshold of −62 dBm for devices with a maximum EIRP of <200 mW and –64dBm for devices with a maximum EIRP of 200 mW to 1 W [12] *averaged over 1 second.*

This is defined as the received signal strength (RSS) (dBm), normalized to the output of a 0dBi receive antenna, that is required to be detected within the WAS channel bandwidth.

2.2 Operational requirements

The WAS should be able to perform channel availability check: A check during which the WAS listens on a particular radio channel for 60 s to identify whether there is a radar operating on that radio channel.

The WAS should be able to perform in-service monitoring: Monitoring of the operating channel to check that a co-channel radar has not moved or started operation within range of the WAS. During in-service monitoring the radar detection function continuously searches for radar signals in-between normal WAS transmissions. This requires the use of quiet spaces between successive WAS transmissions (see Annex 4).

[12] *In practice, it may not be necessary for each device to implement full DFS functionality, provided that such devices are only able to transmit under the control of a device that ensures that all DFS requirements are fulfilled.*

Table B.2 Main parameters of the DFS requirements

Parameter	Value
DFS detection threshold	–62 dBm for devices with a maximum EIRP of <200 mW –64 dBm for devices with a maximum EIRP of ≥200 mW (averaged over one RLAN transmission burst)
Channel availability check time	60 s
Non-occupancy period	30 min
Channel move time	≤10 s

If the WAS has not previously been in operation or has not continuously monitored the channel with in-service monitoring, it should not start transmission in any channel before completion of a channel availability check.

2.3 Response requirements

A channel that has been flagged as containing a radar signal, either by a channel availability check or in-service monitoring, is subject to a 30 min period (non-occupancy period) where it cannot be used by the WAS device in order to protect scanning radars. The non-occupancy period should start at the time when the radar signal is detected.

Additionally, in the band 5 600–5 650 MHz, if a channel has been flagged as containing a radar, a 10 min continuous monitoring of the flagged channel is required prior to use of that channel. Otherwise, other appropriate methods such as channel exclusion would be required.

Channel move time is defined as the period of 10 s needed by a WAS to cease all transmissions on the operating channel upon detection of an interfering signal above the DFS detection threshold. Transmissions during this period will consist of normal traffic for typically less than 100 ms and a maximum of 200 ms after detection of the radar signal. In addition, intermittent management and control signals can be sent during the remaining time to facilitate vacating the operating channel. The aggregate time of the intermittent management and control signals are typically less than 20 ms.

2.4 Summary of the requirements

Table 1 provides a summary of the requirements described above. An example of the operating procedures is given in Annex 2.

Radar signatures were not defined at the time of the WRC-03. Radar signatures were added in the compliance test specifications initially developed by ETSI and later adopted – with changes – by the FCC and Japan's MIC.

The Indoor/Outdoor Issue

Indoor operation attenuates the wireless LAN signal by a large factor – which varies with the building material, proximity to walls or windows, etc. On average, this amounts to 14 dB for large numbers of wireless LANs, affecting the same victim receiver. This is also true of radar systems: due to the high gain antennas, they have a surface coverage that can extend over 10's of square kilometers and, therefore, the aggregate shielding effect is high even though some devices may be used outdoors.

The impact of a small percentage operating close to windows or even outdoors is negligible. The interference potential of handheld devices is limited because of the low power transmitters used.

In practice, the main interference threat comes from "illicit" use of outdoor gear, fixed links using high gain antennas. Illegal frequency settings with or without illegal amplifiers can prove a major source of interference for radars at considerable distances. As the above analysis shows, equipment that meets the latest requirements will adequately protect any radar system. However, technical measures do not necessarily deter intentional illegal use.

The Master/Slave Issue

At the time of WRC 2003, the typical wireless LAN application model was an Access Point serving a number of clients. The wireless LAN industry got the agreement of the regulatory experts at the time to allow the DFS detection and control function to be concentrated in the Access Point and the Access Point controlling the channel of operation so as to avoid interfering with radars. Thus was born the Master/Slave concept: a device could either have the DFS Master role or the DFS Slave role. Based on this distribution of DFS functions, devices operating as DFS slave could roam freely all over the world – the DFS capability of the (static) Access Points would assure that the Slaves would follow the locally applicable rules. Access Points were equated with the DFS Master role and Client devices were equated with the DFS Slave role. In fact, the EU rules require that Slave devices capable of more than 200 mW have their own DFS detection capability – which is nonsensical, given that being a Slave is a role, not an inherent property or type of device.

Since then, the increasing use of peer-to-peer operation and the emergence of My-Fi[13] devices has changed this simple picture, and new ways are needed to clarify device roles and responsibilities in the context of DFS operations.

DFS Compliance: ETSI EN 301–893 and FCC Part 15.407

The work of the ITU-R was used by the European and US Administrations to develop their own DFS certification criteria. These two examples have obtained a wide following: many countries refer to either one or to both as acceptable certification criteria for DFS equipped devices.

[13] Mi-Fy devices combine a 3G cellular capability and a Wi-Fi access point in one package.

ETSI EN 301–893

The ETSI work focused on parameters and test procedures needed to determine if an implementation complies with the requirements. Thus, the EN sits between the regulation embodied in the ECC and EC Directives[14] and the entity responsible for assessing compliance with the regulation. The first operational version of EN 301–893 was version 1.3.1, which provided, in addition to the validation of the DFS threshold, criteria to assess the functional aspect of DFS: the behavior of the device in a "radar environment." Two types of behavior are specified: the (DFS) Master Mode behavior and the Slave Mode behavior. DFS detection requirements only apply to the former and the Slave is assumed to be controlled by the Master, such that the "system" does not cause interference.

The early version had some shortcomings, including the absence of tests for weather radars and not enough randomness in the test criteria – an invitation towards highly specific designs that met the test specifications, but did little else. This was corrected with changes including a minimum pulse width of .8 μs for the whole DFS frequency range, staggered pulse patterns, also for the whole DFS frequency range and a 99.99% detection probability for the weather radar band. Finally, a formula was added for determining the DFS threshold in terms of power spectral density, instead of a power bandwidth independent figure. These changes improve protection of the weather radars, but leave long pulse radars and frequency hopping radars out of scope.

FCC Part 15.407

Encouraged by the wireless industry, the FCC took up the ETSI model of DFS compliance testing; but in discussions with the NTIA, the latter put forward criteria that allow testing the ability of an implementation to detect frequency hopping radars and long pulse (up to 100 μs) radar signals. All detection criteria are given as percentages of detection, given a 50% channel load. The added "NTIA requirements" test radar detection and pulse filtering designs to the limit.

The FCC's DFS requirements include a Channel Availability Check (CAC) before a channel is used, regardless of how long ago the channel has been found free of radar signals. This requirement is not only unnecessary, it is also unfortunate, notably for systems that adapt channel allocations to network load or interference conditions – a device that wants to change channel is blocked for 60 s before it can use a given channel. Given that the CAC is intended to assure fixed radars are detected before a channel is used, the insistence on an "every time CAC" is hard to justify.[15]

[14] ECC Decision (04)08 & EC Decision 2005/513/EC.

[15] ETSI has gone to the other extreme and requires a CAC only at installation time.

The FCC's compliance requirements define "radar waveforms" and detection requirements; the former include variable pulse repetition rates and pulse widths. However, absent were the requirements to protect weather radar systems that use complex scan patterns and staggered pulse patterns. This caused grief when FAA's windshear radars became victims of strong interference in 2007/2008. A modus vivendi was being worked out at the time of writing; it may include[16] a combination of measures such as:

1. exclusion of the 5600–5650 MHz band for all U-NII devices
2. for outdoor equipment only:
   - a 35 km protection zone around their sites, in which the excluded frequency band is modified by the requirement to maintain a frequency separation of 30 MHz relative to the radar's operating frequency
   - professional installation

The above issues point to basic problems with the DFS compliance criteria that hark back to their genesis: a technical problem being solved in a politically complex context. Although different in detail, there are some common shortcomings in "DFS test specifications" of the FCC and ETSI:

1. Insufficient coverage of the radar signal signature space. This may cause interference in future from certified/complaint devices and, therefore, it is a potential threat to the RLAN industry,
2. Overkill in some of the criteria that serve no technically relevant purpose. This reduces the usability or availability of spectrum without reason.
3. Unnecessary complex details increase the threshold of understanding for implementors. This introduces the risk of half-baked implementation.
4. Divergence of criteria between three major markets: US, Europe, and Japan. Technically, there is no reason for these differences because the same chips has to be able to operate in all three markets. Therefore, the criteria could be unified, based on the union of the current sets, suitably modified to correct the above shortcomings. See Chap. 9, Sect. 9.1.4 on the subject of improved DFS compliance criteria.

Behavior Requirements[17] for Master Mode and Slave Mode

The DFS requirements recognize that wireless LANs are used in networks and that some of the network devices control the behavior of other devices, including the choice of operating channel. Therefore, the DFS requirements describe two modes:

[16] In October 2010, no resolution of this issue had been reached.

[17] See Sect. 4.7.1.3 of EN 301–893, V1.5.1.

the Master Mode and the Slave Mode. Typically, the former role is performed by an Access Point and the latter by a client device. Early on, the concepts of role and device were used interchangeably and now the two have become synonymous – which complicates the application of DFS to new modes of use of wireless LAN devices such as "direct link" in which Access Point and Client roles may be switched or performed concurrently.

The behavior required of a device in Master mode is as follows:

1. Scan a channel before use for the duration of the Channel Availability Check time – 10 min in the 5,600–5,650 MHz range, 60 s in the rest of the DFS bands. This is called the Channel Availability Check.

   If no radar signal is detected, the channel may be used immediately; the channel remains available until a radar signal is detected.
2. During the use of a channel, monitor for radar signals – this is called In Service Monitoring.

   If radar signals are detected, then the Master has to stop all data traffic and instruct clients to go off channel – this process may take up to 10 s, during which an aggregated transmission time of 1 s is allowed for channel management exchanges with devices in Slave mode. These periods are known as the Channel Move Time and the Channel Closing Transmission Time.
3. After radar detection, wait at least 30 min before checking and using the same channel again – after performing a Channel Availability Check. This is called the Non-Occupancy Time.

In addition, the ETSI specification describes an "Off Channel CAC": an optional mechanism that allows a device to check another channel than the one it is operating on for the presence of radar signals. The reason for its existence is not clear.

The behavior required of a device in Slave mode is as follows:

4. Do not transmit on a channel before receiving an appropriate enabling signal for a Master device.
5. Stop all transmissions if instructed to do so by the Master [mode] device with which the Slave [mode device] is associated.
6. If operating at more than 200 mW EIRP, do radar detection and stop transmitting if a radar is detected. Perform a Channel Availability Check or Off Channel CAC before using the channel again.

DFS Compliance Criteria

Table B.3 summarizes the key behavioral requirements of ETSI and the FCC. Note that these data are not precise. They are given here only for illustrative purposes. The grayed out rows are unique for the ETSI specification at the time for writing (December 2010).

Table B.3 DFS Behavior compliance criteria and applicability requirements

		Applicability		
Parameter name	Parameter value	Master	Slave with radar detection	Slave without radar detection
Channel availability check time	Outside 5,600–5,650 MHz: 60 s	Yes	Yes	No
	Overlapping with 5,600–5,650 MHz: 600 s	Yes	Yes	No
Maximum off-channel CAC time	Outside 5,600–5,650 MHz: 4 h	Optional	Optional	No
	Overlapping with 5,600–5,650 MHz: 24 h	Optional	Optional	No
DFS detection bandwidth	80% of the RLANs 20 dB bandwidth	Yes	Yes	No
DFS detection threshold	Tx ≤ 200 mW: –62 dBm Tx ≥ 200 mW: – 64 dBm	Yes	Yes	No
DFS detection probability[18]	Outside 5,600–5,650 MHz: ETSI = 60% , FCC = 80%	Yes	Yes	No
	Overlapping with 5,600–5,650 MHz: ETSI 910.99%	Yes	Yes	No
In service monitoring	All frequencies	Yes	Yes	No
Channel move time	10 s	Yes	Yes	Yes
Channel closing transmission time	ETSI: 1 s FCC: 260 ms	Yes	Yes	Yes
Non-occupancy period	30 min	Yes	Yes	No
Uniform spreading	ETSI ≥ 60% of the available frequencies FCC: undefined	Yes	No	No

EN 301–893, Pulse Patterns and Detection Probability

The DFS requirements define a number of pulse patterns that are intended to emulate real-world radars. Version V1.5.1 of the EN defines a number of test patterns that apply across the whole DFS Frequency range. The table below gives only the pulse patterns. Refer to the latest version of the EN for the currently applicable data and other details.

Two properties of these Test Signals are apparent, but hard to explain:

1. Over all Test Signals, a higher PRF is coupled with a higher number of pulses per burst.
2. For a given Test Signal, the number of pulses per burst is constant for large variations in PRF.

[18]This is measured with the RLAN traffic set to 50% of the channel capacity at the highest available data rate.

Table B.4 ETSI, DFS detection test patterns

Radar test signal #	Pulse width [μs] Min	Pulse width [μs] Max	Pulse repetition frequency (PPS) Min	Pulse repetition frequency (PPS) Max	Number of different PRFs	Pulses per burst (per PRF)	Required detection probability
1	0.8	5	200	1,000	1	10 (18)	60%
2	0.8	15	200	1,600	1	15 (18)	60%
3	0.8	15	2,300	4,000	1	25	60%
4 (constant PRF, chirped)	20	30	2,000	4,000	1	20	60%
5 (pulse staggered)	0.8	2	300	400	2/3	10 (18)	60%
6 (pulse staggered)	0.8	2	400	1,200	2/3	15 (18)	60%

The figures in brackets give the number of pulses to be used for testing behavior in the 5,600–5,650 MHz band. Testing in this band has to be done with a signal that exceeds the DFS threshold by +10 dB. This means that compliance with criteria intended to assure the protection of the highly vulnerable weather radar systems is easier to achieve than compliance with the criteria intended to protect other radars – e.g. military radars modeled by Radar Test Signals 1 and 5. On top of that, the pulses per burst apply *per PRF* in case of a staggered PRF test. This too makes the test easy to meet and, therefore, it reduces the *de jure* protection of weather radars. This asymmetry appears to be based on misunderstanding of the subject matter.

The required detection probability is 60%, except for channels overlapping with the 5,600–5,650 MHz range: here the detection probability has to be 99.99%. Confirming this level of probability requires tens of thousands of test runs. This is too time consuming and too costly. Instead, the EN requires a sequence of 20 bursts –- with each burst having to be detected successfully. The effective probability is, therefore, no better than 20 out of 21 = 95.23%, a number that, like the above asymmetry in the criteria, reflects a profound misunderstanding of the subject matter. This too does not add to the *de jure* protection of weather radars. The *de facto* protection is much better because most radar systems have sufficiently high pulse rates to assure reliable detection.

FCC Part 15.407, Pulse Patterns and Detection Probability

Although based on the same source material as ETSI's EN 301–893, the FCC's Part 15.407 requirements as of 2010 are different in some ways:

1. The FCC does not require a 10 min Channel Availability Check for the weather radar sub-band.
2. The FCC does not require a client with more than 200 mW EIRP to perform radar detection; instead, it defines the behavior of an optional configuration: a client device with DFS. This makes sense only if "Client device" is read as "wireless LAN client device." The Closing Channel Transmission time is only 200 ms

Table B.5 FCC: Short Pulse test patterns

Radar test signal #	Pulse width (μs)	Pulse repetition interval (μs)	Number of pulses	Minimum percentage of successful detection	Minimum number of trials
1	1	1,428	18	60%	30
2	1–5	150–230	23–29	60%	30
3	6–10	200–500	16–18	60%	30
4	11–20	200–500	12–16	60%	30
Aggregate (radar types 1–4)				80%	120

upon detection +60 ms aggregate transmission time during the Channel Move Time. These figures stem from the early days of the development of these test specifications. As noted below, DFS detection in most cases occurs in the initial part of the time the radar beam moves over a wireless LAN device. From an interference point of view, a delay of 50–100 ms would be preferable instead of the 200 ms transmission window specified by the FCC. Last but not least, the FCC did not adopt ETSI's approach of a one time Channel Availability Check at installation time. Instead, the FCC requires a Channel Availability Check (CAC) before a channel is used, regardless of how long ago the channel has been found free of radar signals. This requirement is unfortunate for systems that adapt channel allocations to network load or interference conditions – a device is blocked for 60 s before it can use a channel. Given that the check is intended to assure fixed radars are detected before a channel is used, the insistence on an "every time CAC" is hard to justify.

Although the initial work in the US made use of material from ETSI EN 301–893, changes proved inevitable. The NTIA, responsible for coordination of spectrum and telecommunications policy among the agencies of the federal government, put forward criteria that test the ability of a DFS implementation to detect frequency hopping radars and long pulse (up to 100 μs) radar signals. As in the ETSI EN, all detection criteria are given as percentages of detection, given a 50% channel load. The additional "NTIA requirements" test radar detection and pulse filtering designs to the theoretical limit. The following quotes the FCC "radar waveforms" and detection requirements, as these are in force today.[19] Note that these include variable pulse repetition rates and pulse widths. However, absent are the requirements to protect advanced radar systems that use complex scan patterns and staggered pulse patterns. However, given the recent spade of interference cases in the US and the FAA's insistence of a comprehensive solution, changes are inevitable. These will certainly include staggered PRF detection, shorter pulse widths and the requirement for a 10 min Channel Availability Check in the 5,600–5,650 MHz range.

[19] See FCC06-96: Memorandum and Order Revision of Parts 2 and 15 […] U-NII devices […], Appendix).

Table B.6 FCC long Pulse test pattern

Radar type	Pulse width (μs)	Chirp width (MHz)	PRI (μs)	Pulses per burst	Number of bursts	Minimum percentage of successful detection	Minimum number of trials
5	50–100	5–20	1,000–2,000	1–3	8–20	80%	30
6	1	333	9	0.333	300	70%	30

Table B.7 FCC: Frequency Hopping Radar Test Pattern

Radar type	Pulse width (μs)	PRI (μs)	Pulses per hop	Hopping rate (kHz)	Hopping sequence length (ms)	Minimum percentage of detection	Minimum number of trials
6	1	333	9	0.333	300	70%	30

The FCC defines four different "Radar Types" for testing the short pulse detection capability.

Note that the FCC requires an aggregate success rate over all 4 waveforms of 80% – this is much tighter than the 60% of the ETSI EN. The benefit of this tighter requirement is that a compliant detector will perform much better over a broader range of real radar signals.

The FCC also defines a long pulse waveform that emulates a type of radar that makes extensive use of frequency/phase modulated transmission waveforms spread in time. The total number of pulses in the 12 s test period is distributed over 8–20 bursts of equal duration. Each burst may have a different number of pulses. The definition in the FCC's test procedure is a bit more complicated, but the basics are clear: this waveform is very different from ETSI's 30 μs wide pulse test pattern.

Finally, the FCC defines a frequency hopping test pattern – see the table below. The key parameter here is the number of pulses per burst. As explained below in Sect. B.1.4, the nine pulses per burst is at the edge of what is feasible for a detector that sees the channel only 50% of the time.

Taken together with the expected changes to accommodate the testing with pulse patterns related to weather radars, the FCC's "waveforms" offer more assurance of real-world detection performance than those of ETSI.

Appendix C: Radio Resource Management and Wireless Network Management Systems

Radio Resource Management is often a major part of a more comprehensive wireless network management system. There are many wireless network Management Systems, tools and software available on the market, mostly for wireless LAN networks. These tools can be roughly classified as follows:

Site Survey Tools

These tools are often implemented as a third-party tool or a standalone application. There are many off-the-shelf survey tools available. Simple versions only passively observe the frames received and determine the received signal strength and signal quality and other parameters such as station identifiers and traffic loads. More advanced survey tools can actively probe the channels to collect RF data by sending out test frames and evaluating the responses. Technologies for site survey are quite mature, with progressive additions to cover emerging technologies such as 802.11n High Throughput RF signals and devices. Sometimes, GPS technologies are built into these tools to support the mapping device locations and coverage. Site survey tools allow the users to easily collect RF data to generate so called "heat maps" for RF signals or throughput distributions across the covered areas of a deployment. The data collected via such survey tools are highly useful as inputs for Radio Resource Management tools that generate optimal RF plans.

Planning Tools

The scope of Radio Resource Management includes network planning to determine the optimal access point count, optimal access point placements, channel assignment and power assignments. Where there is no real wireless network deployed yet, such planning is usually performed in a predictive way. A typical approach for pre-

dictive Radio Resource Management planning includes importing floor plans for the covered area of the planned deployment, virtually placing access points on the floor plan and computing or simulating the resulting coverage and throughput distributions. A lot of experience and knowledge of the details of RF technologies and wireless technologies are required to come up with a good RF plan. Many trial runs will be required if this is performed manually and, therefore, it can be a very time consuming process. More often than not, RF planning is assisted with some automation of searching for the optimal RF plan. The quality of the resulting RF plan is often hard to assess until it is verified in the real-world deployment.

Depending on how the RF plan is generated, its quality and therefore its effects may vary significantly, ranging from very successful wireless network operation with optimal channel utilization to a very poorly performing network with a lot of coverage holes or pathetic overall system throughput. Therefore, the algorithms for automatic searching for the optimal RF plan are of subtle importance for wireless network planning tools.

Monitoring Tools

Monitoring tools provide the real-time status and statistics for a wireless network. These tools collect a variety of information ranging from simple transmission error rates channel utilization rates and to SNR measurements and histograms of signal strength. The collected information should support analysis methods that determine actual or predicted degradation of link throughput, changes in coverage and other variables that are relevant to the purpose of the network. An advanced monitoring tool will be able to alert the network administrator when any of the above occurs or if the overall network is under-performing. Better yet, it will suggest to the network administrator recommended changes in network or in device configuration. Such features are desirable, but very rare in current offerings. They require ongoing RF management, which often employs an analytic approach to refresh the RF plan with the changes of channel conditions or network conditions.

An analytical Radio Resource Management approach takes statistics from the real world; this reduces the need for familiarity of network administrators with the wireless network and RF technologies. Like the predictive approach, the quality of the generated RF plan depends a lot on the selection of the algorithm, as well as the number of measured variables and how these are taken into account in the algorithm. Applying a poor RF plan to a wireless network can cause devastating degradation of the network's performance and raise major customer complaints. A good algorithm is therefore essential to the analytic approach, just like it is essential for other automation processes.

Both predictive and analytical planning and monitoring tools are readily available. Some of these tools take a hybrid approach and adopt both approaches. They are often built into a full-blown wireless network management system, although standalone tools can also be found.

Wireless Networks and Radio Resource Management Systems

Wireless network management systems are typically much more comprehensive than survey tools, planning tools, and monitoring tools. Although Radio Resource Management can be made a standalone function, it is much more powerful when incorporated into a full-blown wireless network management system. The input statistics collected by the latter's sub-systems, especially when they are collected in real-time, make planning, monitoring, analysis and prediction much more precise and relevant.

Open Platforms are a new trend and the concept has been applied to wireless network management. Open platforms allow sub-systems developed and implemented by a third party to be plugged into the platform. Using an Open Platform based network management systems allows rapid adoption of technical innovation, such as Radio Resource Management, while retaining the more mature or less technically challenging parts such as user interfaces and data collection. Open Platforms also leaves customers more flexibility to choose desirable components that best fit their needs.

Integrated Network and Radio Resource Management

Wireless networks form the edge of the overall network or system; it is usually the part of the network closest to the end users. Likewise, a wireless network management system should be a part of the overall network management system; at least, it should possess the external access interfaces to a network management system. A single point of access for network management makes it easier to view the whole systems, applying network level of policies, as well as applying system level security measures. This aligns with a fundamental perspective that a wireless network should be simply an extension to the existing wired networks solely for the ease of accessibility provided by wireless technologies, nothing more than that.

Radio Resource Management is often integrated into a full-blown wireless LAN management system, although it can be a standalone system as well. A wireless LAN management system usually consists of many sub-systems or modules for various functionalities. Radio Resource Management is among the few essential factors that differentiate one wireless LAN management system from another.

The channel utilization efficiency directly affects the performance of a wireless LAN, which in turn is basically determined by how its Radio Resource Management handles interferences.

Wireless Network Management Services

In addition to the Radio Resource Management sub-system, a wireless network management system usually provides the following set of management services:

Network Monitoring Service

This Service provides a common "*get x*" operation which allows administrators to collect the status and statistics data in real time for the real-world wireless network under its management and showing the status and statistics data at both device level and system level. Such a service usually comes with an advanced graphical user interface that facilitates the understanding of the complex relationships in a wireless network.

Network Control Service

This Service provides a common "*set x*" operation which allows administrators to set configuration parameters and policies at both device level and system level. Such a service usually comes with some graphical user interface. In some advanced wireless network management systems, such "*set x*" operations may be automatic to some extent.

Network Installation Service

The Installation Service for a wireless network management system performs tasks such as device discovery, initial setups, device configuration, policy assignment at both the device level and the system level, either manually or automatically.

Network Reporting, Logging and Trouble Shooting Services

These services provide interfaces to administrators or the end users. They include such functions as status reporting, events logging, alerting of changes in network conditions and channel conditions, and diagnostic information logging and collection. They are essential to the maintenance of a wireless network.

Network Security Management Service[20]

Security management services provide the means for network administrators to install, monitor and control network security services. The security management

[20] Security of wireless networks used to be among the top concerns, but this is no longer the case thanks to efforts in the past few years to adopt a number of industry wireless standards such as Amendments IEEE 802.11i for user security and 802.11w for the security of management frames. Today, the security level for wireless LAN is as good as the wired networks, even from the perspective of enterprise users.

service is a key component of an enterprise wireless enterprise network management system. Security services include setting security policies, device authentication, authorization and upper layer encryption, as well as Network Access Control, which is needed to provide secure web browser access to the wireless Network Management System. Access Control can also include some more advanced security measures, such as Intrusion Detection (IDS) and Intrusion Prevention Systems (IPS), Rogue access point detection, ad-hoc client detection, spoofed SSIDs detection, etc.

Mobility Management Service

Mobility management is important for voice-over-wireless applications, since seamless roaming is required for a satisfactory user experience of phone users. The management of services that support roaming devices, load balancing, and session persistence falls into this category of wireless network management services. Some of them, roaming for example, are supported by industry standards, such as Amendment IEEE 802.11r, which supports fast hand-off has been incorporated into the 2010 version main 802.11 standard.

List of Abbreviations

2 G	Second Generation [Cellular technology]
3 G	Third Generation [Cellular technology]
4 G	Fourth Generation [Cellular technology]
ACR[R]	Adjacent Channel Rejection [Ratio]
ACLR	Adjacent Channel Leakage Ratio
A-IFS	Arbitration Inter-Frame Space
BER	Bit Error Rate
BSS	Basic Service Set [of a wireless LAN network]
CCA	Clear Channel Assessment
CDMA	Code Division Multiple Access
CEPT	Conf. of European Postal and Telecommunication Authorities
CFP	Contention Free Period
CP	Contention Period
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear To Send
CUS	Collective Use of Spectrum [model]
DAA	Detect And Avoid
dBc	decibel relative to carrier [power]
dBi	decibel of inserted loss or gain
dBm	decibel relative 1 milliwatt
dBr	decibel relative to some reference, typically 0
DCF	Distributed Coordination Function [of IEEE 802.11]
DFS	Dynamic Frequency Selection
DIFS	DCF Inter-Frame Space
DSA	Dynamic Spectrum Access
DSSS	Direct Sequence Spread Spectrum
E_b/N_o	Ratio of Energy per bit to Noise power density
EC	European Commission
ECC	European Communications Committee
EDCA	Enhanced Distributed Channel Access [of IEEE 802.11]
EIRP	Equivalent Isotropically Radiated Power
ETSI	European Telecommunications Standards Institute
EY-NPMA	Elimination-Yield-Non-Pre-emptive Multiple Access

FAA	Federal Aviation Agency
FCC	Federal Communications Commission
FDD	Frequency Division Duplex
FHSS	Frequency Hopping Spread Spectrum
GPS	Geographical Positioning System
GSM	General System for Mobile Communications (see also 2 G]
HCCA	HCF Controlled Channel Access
HCF	Hybrid Coordination Function [of IEEE 802.11]
HT	High Throughput [capability of IEEE 802.11]
IEEE	International Institute of Electrical and Electronic Engineers
ITU-R	International Telecommunications Union, Radio [Division]
I/N	Interference to Noise ratio
IoT	Internet of Things
ISM	Industrial, Scientific and Medical [frequency band for…use]
ITS	Intelligent Transportation Systems
LBT	Listen Before Talk
LAN	Local Area Network
LEFR	License Exempt [spectrum] Framework Review
LTE	Long Term Evolution [of 3rd generation cellular technology]
OFDM	Orthogonal Frequency Division Multiplex
OFDMA	OFDM with multiple Access
PAN	Personal Area Network
PRF	Pulse Repetition Frequency
μs	microsecond
ms	millisecond
MAC	Medium Access Control [function or layer]
MCS	Modulation and Coding Scheme [used in IEEE 802.11]
MIMO	Multiple Input, Multiple Output [channel]
MISO	Multiple Input, Single Output [channel]
MLS	Microwave Landing System
MR	Maximum Ratio Combining
NAV	Network [resource] Allocation Vector
NTIA	National Telecommunications and Information Administration
PHY	Physical [layer]
RRM	Radio Resource Management
RTS	Request to Send
SDM	Spatial Division Multiplex
SIMO	Single Input, Multiple Output [channel]
SISO	Single Input, Single Output [channel]
SNR	Signal to Noise Ratio
SNIR	Signal to Noise plus Interference Ratio
SUR	Spectrum Usage Right
TARB	Transmitters as Receiver Beacons [medium access protocol]
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TVBD	TV Band Device
TVWS	TV White Space [spectrum]
UMTS	Universal Mobile Telecommunications System

UWB	Ultra-Wide-Band
V2I	Vehicle to Infrastructure [communications in ITS]
V2V	Vehicle to Vehicle [communications in ITS]
V2X	Vehicle to any [communications in ITS]
WAPECS	Wireless Access Policies for Electronic Communications Systems
WLAN	Wireless LAN
WMAN	Wireless Metropolitan Area Network
WPAN	Wireless Personal Area Network
WRAN	Wireless Regional Area Network
WFA	Wi-Fi Alliance
WRC	World Radio Conference [of the ITU-R]

Bibliography

1. 5 GHz high performance RLAN; Harmonized EN covering the essential requirements of article 3.2 of the R&TTE Directive. Harmonized European Standard (Telecommunications series) EN 301 893 v1.5.1, ETSI BRAN, April 2008
2. ECC Report 120: Technical Requirements for UWB DAA (detect and avoid) devices to ensure the protection of radiolocation services in the bands 3.1–3.4 GHz and 8.5–9 GHz and BWA terminals in the band 3.4–4.2 GHz. June 2008
3. ECC Report 140: Compatibility between RLAN on board Aircraft and Radars in the bands 5250–5350 MHz and 5470–5725 MHz, May 2010
4. ECC Report 159: Technical and operational requirements for the possible operation of cognitive radio systems in the “white spaces” of the frequency band 470–790 MHz
5. Wideband transmission systems; Data transmission equipment operating in the 2,4 GHz ISM band and using wide band modulation techniques;[..]. Harmonized European Standard (Telecommunications series) EN 300 328 v1.7.1, ETSI ERM TG11, October 2006
6. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. ANSI/IEEE Std IEEE Std. 802.11-2007, IEEE, June 2007
7. Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (WPANs). IEEE Std 802.15.4a-2007, IEEE, August 2007
8. N. Abramson, The ALOHA system – Another alternative for computer communications. *Proceedings of the 1970 Fall Joint Computer Conference*, New York, 1970
9. R. Akl, A. Arepally, Dynamic channel assignment in IEEE 802.11 networks. *IEEE International Conference on Portable information Devices, PORTABLE’07*, Orlando, May 2007
10. A.I.F. Akyildiz et al., NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey. Comput. Netw. **50**(13), 2127–2159 (2006)
11. S.M. Alamouti, A simple transmit diversity technique for wireless communications. IEEE J. Sel. Area. Commun. **16**, 1451 (1998)
12. G. Anastasi et al., HIPERLAN/1 MAC protocol: stability and performance analysis. IEEE J. Sel. Area. Commun. **18**(9), 1787–1798 (2000, September)
13. P. Anker, Cognitive radio, the market and the regulator, Dyspan, Singapore (2010)
14. H. Al-Rizzo et al., Enhanced channel assignment and load distribution in IEEE 802.11 WLANs. *IEEE International Conference on Signal Processing and Communications, ICSPC’07*, Dubai, Nov 2007, pp. 768–771
15. C. Andren et al., CCK: the new IEEE 802.11 standard for 2.4 GHZ wireless LANs. *Proceedings of International IC-Taipei Conference*, Taipei, May 2000, pp. 25–39
16. V. Angelakis et al., Adjacent channel interference in 802.11a: Modeling and testbed validation. *IEEE Radio and Wireless Symposium*, pp. 591–594 (Jan 2008)

17. T. Arildsen and F.H.P. Fitzek. *The C-Cube Concept – Combining Cross-Layer Protocol Design, Cognitive, and Cooperative Network Concepts*, ed. by F. Fittzek, M. Katz. Cognitive wireless networks, Chap. 21 (Springer, Dordrecht, 2007)
18. P. Bahl et al., Cell breathing in wireless LANs: Algorithms and evaluation. IEEE Trans. Mob. Comput. **6**(2), 164–178 (2007)
19. M. Barry et al., Distributed control algorithm for service differentiation in wireless frame networks. *Proceedings of IEEE INFOCOM*, Alaska, 2001
20. A. Bazzi et al., Link adaptation algorithms over IEEE 802.11 WLANs in collision prone channels. *IEEE 63 rd Vehicular Technology Conference, VTC'06-Spring*, Melbourne, May 2006
21. Y. Bejerano et al., Fairness and load balancing in wireless LANs using association control. *Proceedings of the 10th International Conference on Mobile Computing and Networking, MobiCom'04*, Philadelphia, 2004, pp. 315–329
22. O. Bendov, Self-interference and loss of SNR within overlapped coverage areas of ATSC On-channel Distributed Transmitters, DTS white Paper, 2008
23. Y. Benker, The commons as a neglected factor of Information Policy, draft, *Telecommunications Policy Research Conference* (1998)
24. L. Berlemann, S. Mangold, *Cognitive Radio* (Wiley, New York, 2009)
25. C. Bianchi, A. Meloni, Natural and man-made terrestrial electromagnetic noise: an outlook. Ann. Geophys. **50**(3), 435–445 (2007)
26. J.W. Betz, Effect of partial-band interference on receiver estimation of C/No, Theory, The MITRE Corporation, 2006: Theory
27. S.Biaz, S. Wu, Rate adaptation algorithms for IEEE 802.11 networks: A survey and comparison. *13th IEEE Symposium on Computers and Communications, ISCC'08*, Marrakech, July 2008
28. B. Bing (ed.), *Emerging Technologies in Wireless LANs* (Cambridge University Press, Cambridge, 2008)
29. D. Br´elaz, New methods to color the vertices of a graph. Commun. ACM **22**, 251–256 (1979)
30. R. Bruno et al., IEEE 802.11 optimal performances: RTS/CTS mechanism vs. basic access. *13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC'02*, vol. 4 September 2002, pp. 1747–1751
31. J. Camp et al., A measurement study of multiplicative overhead effects in wireless networks, in *Mobicom* (2007)
32. P. Chatzimisios, C. Boucouvalas, V. Vistas, Influence of channel BER on IEEE 802.11 DCF. Electron. Lett **39**(23), 1687–1689 (2003)
33. K. Chebrolu et al., Long-distance 802.11b links: Performance measurements and experience. *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*, MobiCom'06, Los Angeles, pp. 74–85 (ACM 2006)
34. Ian Grant, The national broadband plan to set US on path to digital inclusion, Computer Weekly (16 March 2010)
35. T. Chen et al., *A cluster-based cognitive radio network*. 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN'07, Dublin, pp. 168–178 (April 2007)
36. ThM Chen, S.S. Liu, A model and evaluation of distributed network management approaches. IEEE J. Sel. Areas Commun. **20**(4), 850–857 (2002)
37. W.Y. Chen, *Home Networking Basis: Transmission Environments and Wired/Wireless Protocols* (Prentice Hall, Upper Saddle River, 2003)
38. C.F. Chiasserini, R.R. Rao, Coexistence mechanisms for interference mitigation in the 2.4-ghz ism band. IEEE Trans. Wireless Commun. **2(**(5), 964–975 (2003)
39. R. R. Choudhury, N. H. Vaidya, Deafness: A MAC problem in ad hoc networks when using directional antennas, *IEEE International Conference on Network Protocols (ICNP)* Berlin (Oct 2004)
40. R. H. Coase, The federal communications commission*,* 2 Jour. of Law. Econ. **1** (1959)
41. Th Cormen et al., *Introduction to Algorithms*, 3rd edn. (MIT, Cambridge, 2009)

42. D. Costa, On the use of some known methods for T-colourings of graphs. Ann. Oper. Res. **41**, 343–358 (1993)
43. J. del Prado, S. Choi. Link adaptation strategy for IEEE 802.11 WLAN via Received Signal Strength Measurement. *IEEE International Conference on Communications, 2003. ICC '03*, vol. 2, Anchorage (May 2003) pp. 1108–1113
44. W. El-Hajj, H. Alazemi, Optimal frequency assignment for IEEE 802.11 wireless networks. Wireless Commun. Mobile Comput. **9**(1), 131–141 (2009)
45. F. Gamba et al., Towards adaptive wlan frequency management using intelligent agents. *Ad-Hoc, Mobile, and Wireless Networks, Second International Conference, ADHOC-NOW 2003*, Montreal, pp. 116–127 (2003)
46. P. Ferre et al., Throughput analysis of IEEE 802.11 and IEEE 802.11e MAC, *IEEE Wireless Communications and Networking Conference*, Atlanta (March 2004)
47. M. Franceschetti et al., The capacity of wireless networks: Information-theoretic and physical limits. IEEE T. Inform. Theory **55**, 1961–1975 (2009)
48. M. Franceschetti et al., The degrees of freedom of wireless networks: information-theoretic and physical limits, *Allerton Conference* (Sept 2008)
49. E. Garcia et al., Effect of adjacent-channel interference in IEEE 802.11 WLANs. *2nd International conference on cognitive radio Oriented Wireless Networks and Communications, CrownCom'07*, Orlando, pp. 118–125 (Aug 2007)
50. E. Garcia, E. L´opez, R. Vidal, J. Paradells. IEEE Wireless LAN capacity in multicell environments with rate adaptation. *The 18th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC'07*, Athens (Sept 2007)
51. E. Garcia et al., Achievable bandwidth estimation for stations in multi-rate IEEE 802.11 WLAN Cells. *8th IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks, WoWMoM'07*, Helsinky, pp. 1–8 (June 2007)
52. M. Garetto et al., Modeling media access in embedded two flow topologies of multihop wireless networks, *Mobicom* (2005)
53. M. Garetto et al., Modeling per-flow throughput and capturing starvation in CSMA multihop wireless networks. IEEE ACM Trans. Networking **16**(4), 864–877 (2008)
54. M.R. Garey, D.S. Johnson, *Computers and Intractibility: A guide to the theory of NP-Completeness* (W. H. Freeman, New York, 1979)
55. M. Gast, *802.11 Wireless Networks: The Definitive Guide*, 2nd edn. (O'Reilly, Sebastopol, 2005)
56. M. Gerrior, R. W. Woodings, Avoiding interference in the 2.4-GHz ISM band, *Microwave Engineering Europe* (16 Feb 2005)
57. J. Gilbert et al., A 4-GBPS uncompressed wireless HD A/V transceiver chipset. IEEE Micro Mag. **28**(2), 56–64 (2008)
58. R. Gummadi et al., Understanding and mitigating the impact of RF interference on 802.11 networks. *Applications, technologies, architectures, and protocols for computer communications, SIGCOMM'07*, Kyoto, pp. 385–396 (ACM 2007)
59. P. Gupta et al., On capacity scaling in arbitrary wireless networks. IEEE Trans. Inf. Theory **55**(9), 3959–3982 (2009)
60. P. Gupta, P.R. Kumar, The capacity of wireless networks. IEEE Trans. Inf. Theory **46**(2), 388–404 (March 2000)
61. P. Gupta et al., Advances in network information theory, *AMS-DIMACS Series in Discrete Mathematics and Theoretical Computer Science* (Aug 2004)
62. P. Gupta et al., A communication architecture for large heterogeneous wireless networks, *Proceedings Allerton Conference* (Sept 2010)
63. G. Hardin, The tragedy of the commons. Science **162**(3859), 1243–1248 (1968)
64. Hazlett et al., Radio spectrum and the disruptive clarity of coase, George Mason University Research Paper, 10–18 (2010)
65. M. Hata, Empirical formula for propagation loss in land mobile radio services. IEEE Trans. Veh. Technol. **29**(3), 317–325 (1980)
66. S. Haykin, *Communications Systems*, 4th edn. (Wiley, New York, 2000)

67. R. Horvitz, Frequency conversion and new technologies: Challenges and opportunities in spectrum management, *Open Spectrum Foundation*, Serbia (2006)
68. M. Hossain et al., Indoor localization using multiple wireless technologies, *MASS*, Pisa (2007)
69. Z. Huang et al., A busy-tone based directional MAC protocol for ad hoc networks, *Mobicom* (2002)
70. R. Jain, D. Chiu, W. Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared computer systems. DEC Research, Technical Report TR-301 (Sept 1984)
71. T.R. Jensen, B. Toft, *Graph Coloring Problems* (Wiley, New York, 1994)
72. L. Jiang, J. Waalrand, Distributed CSMA for throughput and utility maximization. IEEE ACM Trans. **18**(3), 2010 (2008)
73. Y. de Jong, Measurement and modelling of radiowave propagation in urban microcell*s*, Ph. D. Thesis, TU Eindhoven (2001)
74. Y-J. Kim, Y-J. Suh. An efficient rate switching scheme for IEEE 802.11 wireless LANs. *IEEE 61st Vehicular Technology Conference, VTC'05-Spring* (May 2005)
75. K.K. Leung et al, Outdoor IEEE 802.11 cellular networks: MAC protocol design and performance. *IEEE International Conference on Communications, ICC'02*, Vol. 1 (2002) pp. 595–599
76. W. Klotz. Graph coloring algorithms. Mathematical report, TU-Clausthal (May 2002)
77. P. Kolodzy, Interference temperature: a metric for dynamic spectrum utilization, Int. J. Netw. Manag. (Wiley 2006)
78. D. Kotz et al., The mistaken axioms of wireless-network research. Technical report, Dept. of Computer Science, Dartmouth College (July 2003)
79. D. Kotz et al., Experimental evaluation of wireless simulation assumptions. *Proceedings of the 7th ACM MSWiM*, (2004), pp. 78–82
80. M.Lacage et al., IEEE 802.11 rate adaptation: A practical approach. *Proceedings of the 7th ACM MSWiM*, (Oct 2004), pp. 126–134
81. H. Lee et al., Available bandwidth-based association in IEEE 802.11 wireless LANs. *Proceedings of the 11th international symposium on modeling, analysis and simulation of wireless and mobile systems, MSWiM'08*, (ACM, New York, 2008), pp. 132–139
82. W.C.Y. Lee, *Mobile Communications Design Fundamentals*, 2nd edn. (Wiley, Hoboken, 1993)
83. J. Lee et al., Implementing utility-optimal CSMA, *Communication, Control, and Computing*, Allerton (2009)
84. W. Lemstra et al. (eds.), *The Innovation Journey of Wi-Fi* (Cambridge University Press, Cambridge, 2011)
85. A. Lindgren et al., Evaluation of quality of service schemes for IEEE 802.11 wireless LANs, *Proceedings of IEEE LCN'2001* (2001)
86. X. Ling, K.L. Yeung. Joint access point placement and channel assignment for 802.11 wireless LANs. *IEEE Wireless Communications and Networking Conference 2005. WCNC05*, vol. 3 (March 2005), pp. 1583–1588
87. H. Liu, G. Li, *OFDM-Based Broadband Wireless Networks: Design and Optimization* (Wiley, Malden, 2005)
88. E. Lopez-Aguilera, Contributions to the evaluation and enhancement of WLAN IEEE 802.11 medium access control mechanism. Ph.D. thesis, Technical University of Catalonia (UPC) (June 2008)
89. E. Lopez-Aguilera et al., A study on the influence of transmission errors on WLAN IEEE 802.11 MAC performance. *Wireless Communications and Mobile Computing* (2009)
90. R.K. Mallik, Channel capacity of adaptive transmission with maximal ratio combining in correlated Rayleigh fading. IEEE Trans. Wireless Commun. **3**(4), 1124–1133 (2004)
91. S. Mangold, L. Berlemann, IEEE 802.11k: Improving Confidence in Radio Resource Measurements. *The 16th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC'05*, Berlin, (Sept 2005)

92. K. Medepalli, F.A. Tobagi. Throughput analysis of IEEE 802.11 wireless LANs using an average cycle time approach. *IEEE Global Telecommunications Conference, GLOBECOM'05*, St. Louis, (Nov 2005), pp. 3007–3011
93. A. Mishra et al., Weighted coloring based channel assignment for WLANs. Mob. Comput. Commun. R. **9**(3), 19–31 (2005)
94. A. Mishra et al., Distributed channel management in uncoordinated wireless environments. *Proceedings of the 12th annual international conference on mobile computing and networking, MobiCom'06*, (ACM, New York 2006), pp. 170–181
95. J. Mitola, C. Maguire, Cognitive radio: Making software radios more personal. IEEE Pers. Commun. Mag. **6**, 13–18 (1999)
96. J. Mitola, Cognitive radio: Model based competence for software radios, Licentiate Thesis, KTH, Stockholm (Sept 1999)
97. J. Mitola, Cognitive radio, an integrated agent architecture for software defined radio, Dissertation, KTH (2000)
98. C. Murthy, Cognitive Radio: A (Biased) Overview (2006)
99. M. Neufeld, D. Grunwald, *Deafness and virtual carrier sensing with directional Antenna in 802.11 networks*, Technical Report CU-CS-971-04, University of Colorado At Boulder, May (2004)
100. M.K. Panda et al., Saturation throughput analysis of a system of interfering IEEE 802.11 WLANs. *6th IEEE international symposium on a world of wireless mobile and multimedia networks, WoWMoM'05*, Taormina (June 2005)
101. E. Perahi, R. Stacey, *Next Generation Wireless LANs* (Cambridge University Press, Cambridge, 2008/2010)
102. N. Prasad, A. Prasad (eds.), *WLAN Systems and Wireless IP for Next Generation Communications* (Artech House, co., London, 2002)
103. J. Proakis, M. Salehi, *Digital Communications*, 5th edn. (McGraw Hill, New York, 2008)
104. Q. Li et al., Asynchronous co-channel interference suppression in MIMO OFDM systems, *ICC,* Piscataway (2007)
105. R. Ramanathan, On the performance of Ad hoc networks with beamforming antennas, *ACM MobiHoc,* Long Beach (2001)
106. T.S. Rappaport, *Wireless Communications Principles and Practices*, 2nd edn. (Prentice Hall, Upper Saddle River, 2001)
107. J. Riihijarvi et al., Performance evaluation of automatic channel assignment mechanism for IEEE 802.11 based on graph colouring. *The 17th IEEE international symposium on personal, indoor and mobile radio communications, PIMRC'06*, Helsinki (Sept 2006)
108. J. Riihijarvi et al, Frequency allocation for WLANs using graph colouring techniques. *Second annual conference on wireless on-demand network systems and services, WONS'05*, Washington, DC (Jan 2005), pp. 216–222
109. L.G. Roberts, ALOHA packet system with and without slots and capture. Comp. Commun. Rev. **5**(2), 28–42 (1975)
110. A. Sahai et al., Spectrum sharing, fundamental limits and practical challenges, *Dyspan*, Baltimore (2005)
111. T. Schenk, RF impairments in multiple antenna OFDM, Ph. D. Thesis, TU Eindhoven (2006)
112. J. Shi et al, Measurement and modeling of the origins of starvation in congestion controlled mesh networks, *Proceedings of infocom*, Phoenix (2008)
113. I. Siomina, Radio network planning and resource optimization. Ph.D. thesis, Linkoping University (2007)
114. A. Spyropoulos, C. Raghavendra, Capacity bounds for adhoc networks using directional antennas, *ICC,* Anchorage (2003)
115. Sridharan et al., Implementing backpressure-based rate control in wireless networks, in *Information theory and applications workshop*, San Diego (2009)

116. Sobrinho et al., Why RTS-CTS is not your ideal wireless LAN multiple access protocol, *Proceedings of the IEEE Wireless Communications and Networking Conference 2005*, New Orleans, (March 2005)
117. G. Staple, K. Werbach, The end of spectrum scarcity. IEEE Spectr. Mag. **41**(3), 48–52 (2004)
118. Kim Tae-Suk, Improving spatial reuse through tuning transmit power, carrier sense threshold, and data rate in multihop wireless networks, *Mobicom*, Los Angeles (2006)
119. H.Tanaka, T. Ohira, Beam-steerable planar array antennas using varactor diodes for 60-GHz-band applications, *33rd European microwave Conference*, Munich (2003)
120. R. Tandra, A. Sahai, Noise calibration, delay coherence and SNR walls for signal detection, *IEEE DySpAN,* Berkeley (2008)
121. S. Teotia, Effect of Interchannel Interference on Receiver Performance, Thesis, Virginia Tech (1999)
122. R.W. Thomas et al., Cognitive networks: Adaptation and learning to achieve end-to-end performance objectives. IEEE Commun. Mag. **44**, 51–57 (2006)
123. F.A. Tobagi, L. Kleinrock, Packet switching in radio channels: Part II – the hidden terminal problem in carrier sense multiple access and the busy-tone solution. IEEE Trans. Commun. **23**, 1417–1433 (1975)
124. P. Walke et al., *Wireless Systems* (Wiley, New York, 2006)
125. K. Whitehouse et al., Exploiting the capture effect for collision detection and recovery. *Proceedings of the 2nd IEEE workshop on Embedded Networked Sensors, EmNets'05*, IEEE Computer Society (2005), pp. 45–52
126. S. Xu, T. Saadawi, *Does the IEEE 802.11MAC protocol work well in multi-hop wireless ad hoc networks*, IEEE Comm. Mag. Volume 39 (2001)
127. S. Zeadally et al., Vehicular Ad Hoc networks (VANETS): Status, results, and challenges *Telecommunications Systems* (Dec 2010)
128. A. van Zelst, MIMO OFDM for wireless LANs, Ph. D. thesis, TU Eindhoven (2006)
129. J. Zhu et al., Adaptive CSMA for scalable network capacity in high-density WLAN: A hardware prototyping approach, IEEE Infocom (2006)
130. R.A. Fessenden, "*Wireless Telephony*" Transactions of the American Institute of Electrical Engineers. New York: American Institute of Electrical Engineers (1909)
131. M.R. Garey, D.S. Johnson, *Computers and Intractibility: A guide to the theory of NP-completeness* (1997)
132. Cormen et al., *Introduction to Algorithms, Chap. 16 Greedy Algorithms* p. 329. (1990)

Index

J. Kruys and L. Qian, *Sharing RF Spectrum with Commodity Wireless Technologies: Theory and Practice*, Signals and Communication Technology,
DOI 10.1007/978-94-007-1585-1, © Springer Science+Business Media B.V. 2011